

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2005, 10:01:57 ; Search time 86 Seconds  
(without alignments)  
50.055 Million cell updates/sec

Title: US-09-699-224A-1  
Perfect score: 63  
Sequence: 1 IPVLDEGLFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_23Sep04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	63	100.0	12	4 AAB97116	Aab97116 Peptide m
2	42.5	67.5	15	4 AAB97125	Aab97125 Conserved
3	41	65.1	64	6 ADA35231	Ada35231 Acinetoba
4	41	65.1	232	3 AAY74789	Aay74789 Neisseria
5	41	65.1	232	3 AAY74788	Aay74788 Neisseria
6	41	65.1	691	6 ABU15644	Abu15644 Protein e
7	41	65.1	696	6 ABU22993	Abu22993 Protein e
8	41	65.1	702	6 ABR42508	Abr42508 Coumermyc
9	41	65.1	702	6 ABR42547	Abr42547 Chlorobio
10	41	65.1	724	4 AAU57330	Aau57330 Propionib
11	41	65.1	724	6 ABM53849	Abm53849 Propionib
12	40	63.5	224	4 ABB61313	Abb61313 Drosophil
13	40	63.5	400	6 ABM68527	Abm68527 Photornab
14	40	63.5	441	6 ADA36467	Ada36467 Acinetoba
15	40	63.5	986	4 ABB57810	Abb57810 Drosophil
16	39	61.9	92	3 AAE07086	Aae07086 Human gen
17	39	61.9	93	3 AAG56366	Aag56366 Arabidops
18	39	61.9	100	3 AAG56365	Aag56365 Arabidops
19	39	61.9	135	4 AAE07085	Aae07085 Human gen
20	39	61.9	135	4 AAE07058	Aae07058 Human gen
21	39	61.9	225	4 ABB61361	Abb61361 Drosophil
22	39	61.9	316	4 AAG92033	Aag92033 C glutami
23	39	61.9	676	3 AAB21269	Ab21269 Balanus a
24	39	61.9	688	6 ABU21319	Abu21319 Protein e
25	39	61.9	2893	2 AAW98828	Aaw98828 H. pylori

26	39	61.9	2893	2 AAW71556	Aaw71556 Helicobac
27	39	61.9	2893	6 ABU30749	Abu30749 Protein e
28	38	60.3	51	4 AAB48120	Aab48120 Consensus
29	38	60.3	197	6 ABU47981	Abu47981 Protein e
30	38	60.3	197	7 ADC01060	Adc01060 Enteroha
31	38	60.3	213	5 ABP66209	Abp66209 Bifidobac
32	38	60.3	313	4 AAG82332	Aag82332 S. epider
33	38	60.3	322	6 ABR52942	Abr52942 Protein s
34	38	60.3	322	7 ADK62016	Adk62016 Disease t
35	38	60.3	328	5 ABF38384	Abf38384 Staphyloc
36	38	60.3	483	6 ABU20549	Abu20549 Protein e
37	38	60.3	515	3 AAG31867	Aag31867 Arabidops
38	38	60.3	582	3 AAG31866	Aag31866 Arabidops
39	38	60.3	613	7 ADD27928	Add27928 Human hCO
40	38	60.3	620	3 AAG31865	Aag31865 Arabidops
41	38	60.3	620	5 ABB91842	Abb91842 Herbicida
42	38	60.3	644	3 AAG48152	Aag48152 Arabidops
43	38	60.3	711	3 AAG48151	Aag48151 Arabidops
44	38	60.3	712	2 AAR95648	Aar95648 Thermosta
45	38	60.3	749	3 AAG48150	Aag48150 Arabidops

ALIGNMENTS

RESULT 1

AAB97116  
ID AAB97116 standard; peptide; 12 AA.  
XX  
AC AAB97116;  
XX  
DT 07-AUG-2001 (first entry)  
XX  
DE Peptide mimic #1 of conserved gonococcal mAb 2C7 epitope.  
XX  
KW Peptide mimic; vaccine; gonococcal; epitope; Neisseria gonorrhoeae;  
KW infection; monoclonal antibody 2C7; mAb 2C7.  
XX  
OS Synthetic.  
XX  
PN WC200132692-A2.  
XX  
PD 10-MAY-2001.  
XX  
PF 27-OCT-2000; 2000WO-US029749.  
XX  
PR 29-OCT-1999; 99US-0162491P.  
XX  
PA (RICE/) RICE P A.  
PA (NGAM/) NGAMPASUTADOL J.  
PA (GULA/) GULATI S.  
PI Rice PA, Ngampasutadol J, Gulati S;  
XX WPI; 2001-343473/35.  
DR  
XX New peptide mimics of conserved gonococcal epitopes not present in human  
XX blood group antigens, useful for prophylaxis of Neisseria gonorrhoeae  
XX infections.  
PS Claim 16; Fig 2; 57pp; English.  
XX  
CC The invention relates to novel peptide mimics of conserved gonococcal  
CC epitopes which are not present in human blood group antigens. The peptide  
CC mimics are useful for immunising against Neisseria gonorrhoeae  
CC infections. The present sequence is a peptide mimic which binds to an  
CC epitope of the Neisseria gonorrhoeae monoclonal antibody 2C7. It was  
CC synthesised by random peptide display and, following five rounds of  
CC positive selection with mAb 2C7, it was identified as being able to bind  
XX mAb 2C7 by western blotting  
SQ Sequence 12 AA;

```

Query Match      100.0%; Score 63; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFAP 12
Db 1 IPVLDEGLFAP 12

RESULT 2
AAB97125
ID AAB97125 standard; peptide; 15 AA.
XX
AC AAB97125;
XX
DT 07-AUG-2001 (first entry)
XX
DE Conserved gonococcal mab 2C7 epitope peptide mimic fusion peptide.
XX
KW Peptide mimic; vaccine; gonococcal; epitope; Neisseria gonorrhoeae;
KW infection; monoclonal antibody 2C7; mab 2C7; fusion.
XX
OS Synthetic.
XX
PN WO200132692-A2.
XX
PD 10-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-US029749.
XX
PR 29-OCT-1999; 99US-0162491P.
XX
PA (RICE/) RICE P A.
PA (NGAM/) NGAMPASUTADOL J.
PA (GULA/) GULATI S.
XX
PI Rice PA, Ngampasutadol J, Gulati S;
XX
DR WPI; 2001-343473/36.
XX
PT New peptide mimics of conserved gonococcal epitopes not present in human
PT blood group antigens, useful for prophylaxis of Neisseria gonorrhoeae
PT infections.
XX
PS Claim 23; Fig 2; 57pp; English.
XX
CC The invention relates to novel peptide mimics of conserved gonococcal
CC epitopes which are not present in human blood group antigens. The peptide
CC mimics are useful for immunising against Neisseria gonorrhoeae
CC infections. The present sequence is a peptide mimic which binds to an
CC epitope of the Neisseria gonorrhoeae monoclonal antibody 2C7. Peptides
CC were synthesised by random peptide display and were subjected to five
CC rounds of positive selection with mab 2C7 to identify those able to bind
CC mab 2C7. Two cysteine flanking regions were added to the N- and C-termini
CC of a peptide mimic produced by this method to generate the present
CC sequence, which is also an effective peptide mimic of the mab 2C7 epitope
XX
SQ Sequence 15 AA;

Query Match      67.5%; Score 42.5; DB 4; Length 15;
Best Local Similarity 83.3%; Pred. No. 0.85;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 IPVLDEGLFAP 12
Db 4 IPVL-ENGLFAP 14

RESULT 3
ADA35231
ID ADA35231 standard; protein; 64 AA.
XX
AC ADA35231;

Query Match      65.1%; Score 41; DB 6; Length 64;
Best Local Similarity 80.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VLDENGLFAP 12
Db 36 VLDENGLLKP 45

RESULT 4
AAY74789
ID AAY74789 standard; protein; 232 AA.
XX
AC AAY74789;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 261 protein sequence SEQ ID NO:1052.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX

```

```

PF 30-APR-1999; 99WO-US009346.
XX
XX 01-MAY-1998; 98US-0083758P.
XX 31-JUL-1998; 98US-0094869P.
PR 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0099062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
XX N-PSDB; AAZ53551.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics.
XX
XX Claim 2; Page 605; 1453pp; English.
XX
XX AA253015 to AAZ54536, AAZ54577 to AAZ54615, and AA74253 to AA75941
XX represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of the
XX invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the manufacture
XX of medicaments for treating or preventing infection due to Neisserial
XX bacteria (e.g. meningitis and septicemia), to detect the presence of
XX Neisseria bacteria, or to raise antibodies. They may also be used to
XX screen for agonists or antagonists, which may themselves have use as
XX antibacterial agents. The polynucleotides of the invention may also be
XX used in gene therapy protocols
XX
XX SQ Sequence 232 AA;
XX
XX Query Match 65.1%; Score 41; DB 3; Length 232;
XX Best Local Similarity 64.3%; Pred. No. 41;
XX Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
XX
XX QY 1 IPVLDEGL--FAP 12
XX :|||: ||| |||
XX Db 194 LPVLESNGLDVFAP 207
XX
XX RESULT 5
XX AA74788
XX ID AA74788 standard; protein; 232 AA.
XX
XX AC AA74788;
XX
XX DT 21-MAR-2000 (first entry)
XX
XX DE Neisseria meningitidis ORF 261 protein sequence SEQ ID NO:1050.
XX
XX KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX KW antibacterial; gene therapy.
XX
XX OS Neisseria meningitidis.
XX
XX FN WO9957280-A2.
XX
XX PD 11-NOV-1999.
XX
XX PF 30-APR-1999; 99WO-US009346.
XX

```

```

PR 01-MAY-1998; 98US-0083758P.
PR 31-JUL-1998; 98US-0094869P.
PR 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0099062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
XX (CHIR ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
XX N-PSDB; AAZ53550.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics.
XX
XX Claim 2; Page 604; 1453pp; English.
XX
XX AA253015 to AAZ54536, AAZ54577 to AAZ54615, and AA74253 to AA75941
XX represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of the
XX invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the manufacture
XX of medicaments for treating or preventing infection due to Neisserial
XX bacteria (e.g. meningitis and septicemia), to detect the presence of
XX Neisseria bacteria, or to raise antibodies. They may also be used to
XX screen for agonists or antagonists, which may themselves have use as
XX antibacterial agents. The polynucleotides of the invention may also be
XX used in gene therapy protocols
XX
XX SQ Sequence 232 AA;
XX
XX Query Match 65.1%; Score 41; DB 3; Length 232;
XX Best Local Similarity 64.3%; Pred. No. 41;
XX Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
XX
XX QY 1 IPVLDEGL--FAP 12
XX :|||: ||| |||
XX Db 194 LPVLESNGLDVFAP 207
XX
XX RESULT 6
XX ABU19644
XX ID ABU19644 standard; protein; 691 AA.
XX
XX AC ABU19644;
XX
XX DT 19-JUN-2003 (first entry)
XX
XX DE Protein encoded by Prokaryotic essential gene #5171.
XX
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX OS Borrelia cepacia.
XX
XX FN WO20027183-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX

```

```
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA23514.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 47568; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation or the activity of a gene in an operon required for
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 591 AA;
XX
XX Query Match 65.1%; Score 41; DB 6; Length 691;
XX Best Local Similarity 88.9%; Pred. No. 1.5e+02;
XX Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 IPVLDENGL 9
XX |||||:|
XX 674 IPVLDENGL 682
XX
XX RESULT 7
XX ABU22993
XX ID ABU22993 standard; protein; 696 AA.
XX AC ABU22993;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #8520.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Bordetella pertussis.
XX
```

```
WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WPI; 2003-029926/02.
N-PSDB; ACA6863.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 50917; 1766pp; English.
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation or the activity of a gene in an operon required for
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than S. aureus, S. typhimurium,
K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
Sequence 696 AA;
Query Match 65.1%; Score 41; DB 6; Length 696;
Best Local Similarity 72.7%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 IPVLDENGLFA 11
Db :|||:|
682 FPFVLDSDGLXA 692
RESULT 8
ABR42508
ID ABR42508 standard; protein; 702 AA.
XX
```

```

AC ABR42508;
DT 26-AUG-2003 (first entry)
DE Coumermycin A1 resistance gene.
KW Coumermycin A1; aminocoumarin; antibiotic; antibacterial; cytostatic;
KW resistance.
XX Streptomyces rishiriensis.
OS WO2003014352-A2.
XX 20-FEB-2003.
XX 06-AUG-2002; 2002WO-EP008777.
XX 08-AUG-2001; 2001US-0310808P.
XX (UYTU-) UNIV TUEBINGEN.
PA Heide L, Li S;
XX WPI: 2003-268202/26.
XX N-PSDB; ACC58252.
XX Novel polypeptide encoded by open reading frame of aminocoumarin
PT biosynthetic gene cluster, useful for synthesizing aminocoumarins useful
PT for treating bacterial infections and malignancies.
XX Example 1; Fig 1C; 124pp; English.
XX The present sequence is the protein sequence of a resistance protein
CC encoded by the gyrB gene of a newly isolated coumermycin A1 biosynthetic
CC gene cluster of Streptomyces rishiriensis DSM 40489. The sequence shows
CC similarity to Streptomyces coelicolor DNA gyrase subunit B. The
CC coumermycin A1 gene cluster shows marked sequence similarity to the
CC novobiocin gene cluster. The invention provides aminocoumarin (e.g.
CC coumermycin A1) biosynthetic nucleic acids and polypeptides, a method for
CC producing hybrid antibiotics using a combination of various regions of
CC biosynthetic gene clusters, and a pharmaceutical composition containing
CC an aminocoumarin compound for use in the treatment of Gram-positive
CC bacterial infections and of malignant diseases
XX
SQ Sequence 702 AA;
Query Match 65.1%; Score 41; DB 6; Length 702;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLFAP 12
Db 285 VPVLDEHGQWTP 296
:|||||:|

RESULT 9
ABR42547
ID ABR42547 standard; protein; 702 AA.
AC ABR42547;
XX 23-OCT-2003 (revised)
DT 26-AUG-2003 (first entry)
XX Clorobiocin resistant topoisomerase II subunit B.
XX Clorobiocin; aminocoumarin; antibiotic; antibacterial; cytostatic;
KW topoisomerase; enzyme.
XX Streptomyces roseochromogenes subsp. oscitans.
OS WO2003014352-A2.
PN
XX

```

```

PD 20-FEB-2003.
XX 06-AUG-2002; 2002WO-EP008777.
XX 08-AUG-2001; 2001US-0310808P.
XX (UYTU-) UNIV TUEBINGEN.
PA Heide L, Li S;
XX WPI: 2003-268202/26.
XX N-PSDB; ACC58254.
XX Novel polypeptide encoded by open reading frame of aminocoumarin
PT biosynthetic gene cluster, useful for synthesizing aminocoumarins useful
PT for treating bacterial infections and malignancies.
XX Example 2; Fig 2B; 124pp; English.
XX The present sequence is the clorobiocin resistant topoisomerase II
CC subunit B encoded by an open reading frame (ORF) of the clorobiocin
CC biosynthetic gene cluster of Streptomyces roseochromogenes var. oscitans
CC DS 12.976. The clorobiocin biosynthetic gene cluster includes 27 ORFs
CC that show similarity to the corresponding ORFs of the novobiocin and
CC coumermycin A1 biosynthetic gene clusters, and in identical order. The
CC invention provides aminocoumarin (e.g. clorobiocin) biosynthetic nucleic
CC acids and polypeptides, a method for producing hybrid antibiotics using a
CC combination of various regions of biosynthetic gene clusters, and a
CC pharmaceutical composition containing an aminocoumarin compound for use
CC in the treatment of Gram-positive bacterial infections and of malignant
CC diseases. (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ Sequence 702 AA;
Query Match 65.1%; Score 41; DB 6; Length 702;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLFAP 12
Db 286 VPVLDEHGQWTP 297
:|||||:|

RESULT 10
AAU57330
ID AAU57330 standard; protein; 724 AA.
AC AAU57330;
XX 27-FEB-2002 (first entry)
DT Propionibacterium acnes immunogenic protein #18226.
XX DE
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX OS
XX Propionibacterium acnes.
XX WO200181581-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US012865.
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX (CORI-) CORIXA CORP.
PA
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

```

```

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
DR N-PSDB; AAS95982.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 18525; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 724 AA;
XX
XX Query Match 65.1%; Score 41; DB 4; Length 724;
XX Best Local Similarity 58.3%; Pred. No. 1.6e+02;
XX Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 IPVLDENGLFAP 12
Db :|:|||||
309 VPMLDENGGMTP 320

RESULT 11
ID ABM53849
AC ABM53849 standard; protein; 724 AA.
XX
XX ABM53849;
XX
XX 20-OCT-2003 (first entry)
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #18525.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
XX
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skelky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Carter D;
XX Barth B, Vallieve-Douglass J;
XX WPI; 2003-381789/36.
XX

```

---

```

DR N-PSDB; ACF64511.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 18525; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention, antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 724 AA;
XX
XX Query Match 65.1%; Score 41; DB 6; Length 724;
XX Best Local Similarity 58.3%; Pred. No. 1.6e+02;
XX Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 IPVLDENGLFAP 12
Db :|:|||||
309 VPMLDENGGMTP 320

RESULT 12
ID ABB61313
XX
XX ABB61313 standard; protein; 224 AA.
XX
XX ABB61313;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 10731.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX

```

PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 DR N-PSDB; ABL05416.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 10731; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABBS72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 224 AA;  
 Query Match 63.5%; Score 40; DB 4; Length 224;  
 Best Local Similarity 60.0%; Pred. No. 60;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 IPVLDNGGLF 10  
 Db :||:|:|:  
 57 VPMDDNGTGF 66  
 RESULT 13  
 ABM68527  
 ID ABM68527 standard; protein; 400 AA.  
 XX  
 AC ABM68527;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Photorhabdus luminescens protein sequence #1624.  
 XX  
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 KW detection; food; gene expression; plant; animal; microorganism; toxin;  
 KW antibiotic; biopesticide; virulence factor; disease model; plague;  
 KW whooping cough.  
 XX  
 OS Photorhabdus luminescens.  
 XX  
 PN WO200294867-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 07-FEB-2002; 2002WO-IB003040.  
 XX  
 PR 07-FEB-2001; 2001FR-00001659.  
 XX  
 XX (INSP ) INST PASTEUR.  
 XX (CNRS ) CNRS CENT NAT RECH SCI.  
 PA  
 PA Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
 PI Buchrieser C;  
 XX  
 DR WPI; 2003-148459/14.  
 XX  
 CC Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
 XX  
 PS Claim 2; SEQ ID NO 1624; 1205pp; French.  
 XX  
 CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are

CC sources of probes and primers for detecting the genome of P. luminescens  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than P. luminescens and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by P.  
 CC luminescens. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which P.  
 CC luminescens is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated P. luminescens proteins  
 XX  
 SQ Sequence 400 AA;  
 Query Match 63.5%; Score 40; DB 6; Length 400;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 IPVLDNGGLPAP 12  
 Db :||:|:|:  
 12 LPAIDKNGINAP 23  
 RESULT 14  
 ADA36467  
 ID ADA36467 standard; protein; 441 AA.  
 XX  
 AC ADA36467;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Acinetobacter baumannii protein #3628.  
 XX  
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
 KW plant biocontrol agent.  
 XX  
 OS Acinetobacter baumannii.  
 XX  
 PN US6562958-B1.  
 XX  
 PD 13-MAY-2003.  
 XX  
 PF 04-JUN-1999; 99US-00328352.  
 XX  
 PR 09-JUN-1998; 98US-0088701P.  
 XX  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton G, Bush D;  
 XX  
 DR WPI; 2003-576092/54.  
 DR N-PSDB; ADA32341.  
 XX  
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
 PT for diagnosing a bacterial disease, as components of antibacterial  
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
 PT plants.  
 XX  
 PS Example; SEQ ID NO 7754; 328pp; English.  
 XX  
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.  
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
 CC for diagnosing a bacterial disease, as components of antibacterial  
 CC vaccines, as targets for antibacterial drugs, to detect the presence of  
 CC A. baumannii and other Acinetobacter species in a sample, in screening

CC compounds for the ability to interfere with the A. baumannii life cycle  
 CC or to inhibit A. baumannii infection, and as biocontrol agents for  
 CC plants. The present sequence represents the amino acid sequence of an A.  
 CC baumannii protein.

XX Sequence 441 AA;

Query Match 63.5%; Score 40; DB 6; Length 441;  
 Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDENGLEAP 12

Db 118 LPTYDQVGLFAP 129

RESULT is

ABBS7810  
 ID ABBS7810 standard; protein; 986 AA.

XX AC ABBS7810;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 222.

XX KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL01913.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX FS Disclosure; SEQ ID NO 222; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABBS72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 986 AA;

Query Match 63.5%; Score 40; DB 4; Length 986;  
 Best Local Similarity 70.0%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDENGLF 10

Db 297 IPILSEDGLF 306

RESULT 16  
 AAE07086

ID AAE07086 standard; protein; 92 AA.

XX AC AAE07086;

XX DT 16-OCT-2001 (first entry)

XX DE Human gene 8 encoded secreted protein HHTLB76, SEQ ID NO:103.

XX Human secreted protein; proliferative disorder; cancer; tumour;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;  
 KW gastrointestinal disorder; pregnancy-related disorder;  
 KW endocrine disorder; infection; wound healing; vulnerability; cell culture;  
 KW chemotaxis; food additive; gene therapy; binding partner identification.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..24

FT /label= Signal\_peptide

FT Protein 25..92

FT /label= Mature\_human\_secreted\_protein

FT Misc-difference 68

FT /label= Unknown

FT /note= "Encoded by GYT"

FT Misc-difference 86

FT /label= Unknown

FT /note= "Encoded by CAN"

FT Misc-difference 90

FT /label= Unknown

FT /note= "Encoded by NAC"

FT Misc-difference 92

FT /label= Unknown

FT /note= "Encoded by ANA"

XX WO200154708-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001434.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 18-AUG-2000; 2000US-0226279P.

XX PR 05-DEC-2000; 2000US-0251988P.

XX PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PA Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;

XX PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;

XX PI Fiscella M, Ni J, Ruben SM, Barash SC;

XX WPI; 2001-488743/53.

XX DR N-PSDB; AAD13380.

XX New isolated nucleic acids and polypeptides, useful for diagnosing,  
 PT treating and/or preventing human diseases and disorders.

XX Claim 11; Page 521-522; 558pp; English.

XX AAD13345-AAD13401 represent cDNAs corresponding to 22 human secreted  
 CC protein genes, and AAE07051-AAE07105 represent the proteins they encode.  
 CC AAE07106-AAE07129 represent human secreted protein fragments or variants.  
 CC The genes and their secreted proteins are useful for preventing, treating

CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 22 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The  
 CC present sequence represents a human secreted protein of the invention  
 XX  
 SQ Sequence 92 AA;

Query Match 61.9%; Score 39; DB 4; Length 92;

Best Local Similarity 63.6%; Pred. No. 32;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PVLDENGLPAP 12

Db 46 PLIQETGAFAP 56

# RESULT 17

AAG56366  
 ID AAG56366 standard; protein; 93 AA.

AC AAG56366;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 72444.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128214P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130819P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0132456P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 18-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 24-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.

```
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147933P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158222P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159223P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159310P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 61.9%; Score 39; DB 3; Length 93;
Best Local Similarity 45.5%; Pred. No. 33;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFA 11
Db 44 LPIVEENGIFS 54
:|: :|||:|:

RESULT 18
AAG56365
ID AAG56365 standard; protein; 100 AA.
XX AC AAG56365;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 72443.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 03-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 06-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.
```

PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142330P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147460P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147915P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149920P.  
PR 23-AUG-1999; 99US-0150866P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158432P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.

```

PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 61.9%; Score 39; DB 3; Length 100;
Best Local Similarity 45.5%; Pred. No. 36;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDENGLEA 11
Db 51 LPIVEENGIFS 61

RESULT 19
AAE07085
ID AAE07085 standard; protein; 135 AA.
AC AAE07085;
XX
DT 16-OCT-2001 (first entry)
DE Human gene 8 encoded secreted protein HHTLB76, SEQ ID NO:102.
XX Human secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
KW chemotaxis; food additive; gene therapy; binding partner identification.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT Protein 25..135
FT Misc-difference 125
FT /label= Mature_human_secreted_protein
FT /label= Unknown
FT /note= "Encoded by GYG"
XX
PN WO200154708-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001434.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 18-AUG-2000; 2000US-0246279P.
XX
XX 05-DEC-2000; 2000US-0251988P.
XX
XX 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;
XX Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
XX Fiacella M, Ni J, Ruben SM, Baraah SC;
XX WPI; 2001-488743/53.

```

---

```

DR N-PSDB; AAD13379.
XX
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
XX treating and/or preventing human diseases and disorders.
XX
XX Claim 11; Page 521; 558pp; English.
XX
XX AAD13345-AAD13401 represent cDNAs corresponding to 22 human secreted
XX protein genes, and AAE07051-AAE07105 represent the proteins they encode.
XX AAE07106-AAE07129 represent human secreted protein fragments or variants.
XX The genes and their secreted proteins are useful for preventing, treating
XX or ameliorating medical conditions, e.g., by protein or gene therapy.
XX Pathological conditions can be diagnosed by determining the amount of the
XX new protein in a sample or by determining the presence of mutations in
XX the new genes. Specific uses are described for each of the 22 genes,
XX based on the tissues in which they are most highly expressed, and include
XX developing products for the diagnosis or treatment of proliferative
XX disorders, cancer, tumours, foetal and developmental abnormalities,
XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX diseases (e.g., rheumatoid arthritis), inflammation, allergies,
XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
XX angiogenic disorders, kidney disorders, gastrointestinal disorders,
XX pregnancy-related disorders, endocrine disorders, and infections. The
XX proteins can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, for supporting cell culture of primary tissues,
XX to regenerate tissues, to identify their cognate ligands or binding
XX partners, and in chemotaxis, and can be used as a food additive or
XX preservative to modify storage properties. Antibodies specific for a
XX protein of the invention can be used in alleviating symptoms associated
XX with the disorders mentioned above, and in diagnostic immunoassays e.g.,
XX radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
XX present sequence represents a human secreted protein of the invention
XX
XX Sequence 135 AA;
SQ

```

---

```

Qy 2 PVLDENGLFAP 12
Db 46 PLLQETGAFAP 56

Query Match 61.9%; Score 39; DB 4; Length 135;
Best Local Similarity 63.6%; Pred. No. 51;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 20
AAE07058
ID AAE07058 standard; protein; 135 AA.
XX
XX AAE07058;
XX
XX 16-OCT-2001 (first entry)
XX
XX Human gene 8 encoded secreted protein HHTLB76, SEQ ID NO:75.
XX
XX Human secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulnerability; cell culture;
XX chemotaxis; food additive; gene therapy; binding partner identification.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..24
XX Protein 25..135
XX Misc-difference 125
XX /label= Unknown
XX /note= "Encoded by GYG"
XX
XX WO200154708-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001434.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 18-AUG-2000; 2000US-0246279P.
XX
XX 05-DEC-2000; 2000US-0251988P.
XX
XX 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;
XX Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
XX Fiacella M, Ni J, Ruben SM, Baraah SC;
XX WPI; 2001-488743/53.

```

FT Protein 25. .135  
 XX /label= Mature\_human\_secreted\_protein  
 PN WO200154708-A1.  
 XX 02-AUG-2001.  
 PD  
 XX 17-JAN-2001; 2001WO-US001434.  
 PF  
 XX 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;  
 PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;  
 PI Fiacella M, Ni J, Ruben SM, Barash SC;  
 XX  
 DR WPI; 2001-488743/53.  
 DR N-PSDB; AAD13352.  
 XX  
 XX New isolated nucleic acids and polypeptides, useful for diagnosing,  
 PT treating and/or preventing human diseases and disorders.  
 PT  
 XX Claim 11; Page 494; 558pp; English.  
 PS  
 XX AAD13345-AAD13401 represent cDNAs corresponding to 22 human secreted  
 CC protein genes, and AAE07051-AAE07105 represent the proteins they encode.  
 CC AAE07106-AAE07129 represent human secreted protein fragments or variants.  
 CC The genes and their secreted proteins are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 22 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The  
 CC present sequence represents a human secreted protein of the invention  
 XX  
 SQ Sequence 135 AA;  
 Query Match 61.9%; Score 39; DB 4; Length 135;  
 Best Local Similarity 63.6%; Pred. No. 51;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 OY 2 PVLDENGLFAP 12  
 Db :|:|:|:|:|:|  
 46 PLLQETGAFAP 56  
 RESULT 21  
 ABB61361  
 ID ABB61361 standard; protein; 225 AA.  
 XX

AC ABB61361;  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX Drosophila melanogaster polypeptide SEQ ID NO 10875.  
 DE  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEXE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 DR  
 DR N-PSDB; ABL05464.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 XX Disclosure; SEQ ID NO 10875; 21pp + Sequence Listing; English.  
 PS  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 225 AA;  
 Query Match 61.9%; Score 39; DB 4; Length 225;  
 Best Local Similarity 87.5%; Pred. No. 93;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 IPVLDENG 8  
 Db :|:|:|:|:|:|  
 56 IPVLDENG 63  
 RESULT 22  
 AAG92033  
 ID AAG92033 standard; protein; 316 AA.  
 XX  
 XX AAG92033;  
 AC  
 XX 26-SEP-2001 (first entry)  
 DT  
 XX C glutamicum protein fragment SEQ ID NO: 5787.  
 DE  
 XX Corynebacterium; amino acid synthesis; vitamin; saccharide;  
 KW organic acid synthesis.  
 KW  
 XX Corynebacterium glutamicum.  
 OS  
 XX EPI108790-A2.  
 PN  
 XX 20-JUN-2001.  
 PD

XX 18-DEC-2000; 2000EP-00127688.  
 XX  
 XX 16-DEC-1999; 99JP-00377484.  
 PR 07-APR-2000; 2000JP-00159162.  
 PR 03-AUG-2000; 2000JP-00280988.  
 XX  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 PA  
 XX Nakagawa S, Mizoguchi H, Ando S, Hayaashi M, Ochiai K, Yokoi H;  
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 XX WPI; 2001-376931/40.  
 DR N-PSDB; AAH67252.  
 XX  
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analyzing  
 PT expression profile or pattern of a gene and identifying homologous gene.  
 XX  
 PS Claim 17; SEQ ID NO 5787; 246pp + Sequence Listing; English.  
 XX  
 XX The present invention provides a number of nucleotide and protein  
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of coryneform bacterium, measuring expression amount and analysing  
 CC the expression profile or expression pattern of a gene derived from  
 CC Coryneform bacterium, and identifying a homologue of a gene derived from  
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino  
 CC acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a protein described in the  
 CC exemplification of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from the European Patent Office  
 XX  
 XX Sequence 316 AA;  
 SQ  
 Query Match 61.9%; Score 39; DB 4; Length 316;  
 Best Local Similarity 77.8%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 LDENGLFAP 12  
 Db :|||||  
 167 VDENGKFP 175  
 RESULT 23  
 ID AAB23269 standard; protein; 676 AA.  
 XX  
 AC AAB23269;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 XX Balanus amphitrite adhesion/metamorphosis-related protein Bcs-3.  
 DE  
 XX Adhesion/metamorphosis-related protein Bcs-3; barnacle; larva-specific;  
 KW adhesion inhibition; metamorphosis inhibition; compound screening;  
 KW antifouling composition.  
 XX  
 OS Balanus amphitrite.  
 XX  
 XX JP2000228985-A.  
 PN  
 XX 22-AUG-2000.  
 PD  
 XX 09-FEB-1999; 99JP-00031067.  
 PF  
 XX 09-FEB-1999; 99JP-00031067.  
 PR  
 XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.  
 PA  
 XX WPI; 2000-649634/63.  
 DR  
 DR N-PSDB; AAA97603.  
 XX

XX Novel barnacle kipris larva-specific adhesion/metamorphosis-related gene  
 PT useful for identifying inhibitors of barnacle adhesion.  
 XX  
 XX Claim 3; Page 15-17; 32pp, Japanese.  
 XX  
 CC The invention relates to six larva-specific adhesion/metamorphosis-  
 CC related genes from the barnacle Balanus amphitrite (cDNAs given in  
 CC AAA97601-A97606) and to the proteins they encode (AAB23267-B23272). The  
 CC genes and the proteins can be used for screening for a substance that  
 CC inhibits the adhesion or metamorphosis of barnacle larvae, which may be  
 CC useful in antifouling compositions for use in the shipping industry. The  
 CC present sequence represents the Balanus amphitrite adhesion/metamorphosis  
 CC -related protein Bcs-3  
 XX  
 SQ Sequence 676 AA;  
 Query Match 61.9%; Score 39; DB 3; Length 676;  
 Best Local Similarity 87.5%; Pred. No. 3.4e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 DENGFLAP 12  
 Db :|||||  
 563 DENGFLFP 570  
 RESULT 24  
 ID ABU21319 standard; protein; 688 AA.  
 XX  
 AC ABU21319;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 XX Protein encoded by Prokaryotic essential gene #6846.  
 DE  
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 KW  
 XX Burkholderia fungorum.  
 OS  
 XX WO200277183-A2.  
 PN  
 XX 03-OCT-2002.  
 PD  
 XX 21-MAR-2002; 2002WO-US009107.  
 PF  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR  
 DR N-PSDB; ACA25189.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 25; SEQ ID NO 49243; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide(s) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC on a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 688 AA;

Query Match 61.9%; Score 39; DB 6; Length 688;  
 Best Local Similarity 77.8%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENGL 9  
 |||||:  
 Db 672 IPVLDEGDM 680

RESULT 25  
 AAW98828  
 ID AAW98828 standard; protein; 2893 AA.

XX AAW98828;  
 XX  
 DT 31-MAR-1999 (first entry)  
 XX  
 DE H. pylori GHPO 1484 protein.  
 XX  
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.  
 XX Helicobacter pylori.  
 OS  
 PN WO9843478-A1.  
 XX  
 PD 08-OCT-1998.  
 XX  
 PF 01-APR-1998; 98WO-US006371.  
 XX  
 PR 01-APR-1997; 97US-00833457.  
 PR 24-JUN-1997; 97US-00881227.  
 PR 29-JUL-1997; 97US-00902615.  
 XX  
 PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;  
 XX  
 DR WPI; 1998-542293/46.  
 DR N-PSDB; AAX14547.  
 XX  
 PT New isolated Helicobacter polynucleotides - used to develop products for  
 PT the diagnosis, prevention and treatment of Helicobacter infections and  
 PT gastrointestinal diseases.  
 XX

PS Claim 8; Page 1827-1840; 2054pp; English.  
 CC This sequence represents a Helicobacter pylori GHPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastroduodenal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be used for  
 CC detection and diagnosis  
 XX  
 XX Sequence 2893 AA;  
 SQ  
 Query Match 61.9%; Score 39; DB 2; Length 2893;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+03;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLPAP 12  
 |||||:  
 Db 2077 IPNLGKKGLFAP 2088

RESULT 26  
 AAW71556  
 ID AAW71556 standard; protein; 2893 AA.

XX AAW71556;  
 XX  
 DT 09-NOV-1998 (first entry)  
 XX  
 DE Helicobacter polypeptide GHPO 1484.  
 XX  
 KW GHPO 1484; infection; therapy; diagnosis; vaccine; gastritis; ulcer.  
 XX Helicobacter pylori.  
 OS  
 PN WO9821225-A1.  
 XX  
 PD 22-MAY-1998.  
 XX  
 PF 14-NOV-1997; 97WO-US021353.  
 XX  
 PR 14-NOV-1996; 96US-00749051.  
 PR 01-APR-1997; 97US-00831309.  
 PR 01-APR-1997; 97US-00833457.  
 PR 01-APR-1997; 97US-00834705.  
 PR 24-JUN-1997; 97US-00881227.  
 PR 29-JUL-1997; 97US-00902615.  
 XX  
 PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Haas R, Kleanthous H, Tomb J, Miller C, Al-Garawi A, Odenbreit S;  
 PI Meyer T;  
 XX  
 DR WPI; 1998-297855/26.  
 DR N-PSDB; AAV52091.  
 XX  
 PT Helicobacter polynucleotide and polypeptide sequences - useful to treat  
 PT or prevent gastrointestinal infection.  
 XX  
 PS Claim 1; Page 330-337; 362pp; English.  
 XX  
 CC This claimed Helicobacter pylori polypeptide, designated GHPO 1484, can  
 CC be used in vaccination methods for preventing or treating Helicobacter  
 CC infection. 85 Helicobacter polypeptides (see AAW71474-W71558) are  
 CC claimed, as well as isolated polynucleotides (see AAV52009-93) that  
 CC encode them. The invention also provides: methods for producing these  
 CC Helicobacter polypeptides in recombinant host systems, and related  
 CC expression cassettes, vectors and transformed host cells;  
 CC live vaccine vectors that contain the polynucleotides of the invention  
 CC and which can be used to prevent or treat Helicobacter infection;  
 CC therapeutic and/or prophylactic methods involving administration of

CC polynucleotide molecules, polypeptides or monospecific antibodies;  
 CC methods for detecting the presence of *Helicobacter* in samples using e.g.  
 CC the polypeptides or monospecific antibodies; and methods for purifying  
 CC the polypeptides by antibody-based affinity chromatography  
 XX  
 SQ Sequence 2893 AA;  
 Query Match 61.9%; Score 39; DB 2; Length 2893;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+03;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 IPVLDENGLFAP 12  
 ||| : ||||  
 DB 2077 IPNLGKGLFAP 2088  
 RESULT 27  
 ABU30749  
 ID ABU30749 standard; protein; 2893 AA.  
 XX  
 AC ABU30749;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #16276.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS *Helicobacter pylori*.  
 XX  
 FN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX  
 FR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Tarwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA34619.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 25; SEQ ID NO 58673; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 2893 AA;  
 Query Match 61.9%; Score 39; DB 6; Length 2893;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+03;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 IPVLDENGLFAP 12  
 ||| : ||||  
 DB 2077 IPNLGKGLFAP 2088  
 RESULT 28  
 AAB48120  
 ID AAB48120 standard; peptide; 51 AA.  
 XX  
 AC AAB48120;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE Consensus thyroglobulin type 1 domain.  
 XX  
 KW TANGO 204; TANGO 206; TANGO 209; A236; secreted protein; human; mouse;  
 KW transmembrane protein; antianemic; cerebroprotective; arteriosclerosis;  
 KW antiaesthmic; neuroprotective, cytostatic; cardiant; hepatotropic;  
 KW antiinflammatory; antidiabetic; antifertility; antipyretic; vasotropic;  
 KW antirheumatic; nephrotropic; hemostatic; antilipemic; osteopathic;  
 KW ophthalmologic; antisickling; antiulcer; vulnerary; thyroglobulin.  
 XX  
 OS Unidentified.  
 XX  
 XX WO200069885-A2.  
 XX  
 XX 23-NOV-2000.  
 XX  
 XX 15-MAY-2000; 2000WO-US013361.  
 XX  
 XX 14-MAY-1999; 99US-00312359.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Pan Y, Leiby KR;  
 PI  
 XX WPI; 2001-024999/03.  
 DR  
 XX Novel nucleic acids encoding secreted or transmembrane proteins, useful  
 PT for treating, e.g. cancer, hemophilia, anemia, ischemia or diseases of  
 PT the lung, liver, kidney or pancreas.  
 XX  
 XX Disclosure; Fig 16; 209pp; English.  
 XX  
 CC The invention provides human and mouse nucleic acids designated TANGO  
 CC 204, TANGO 206, TANGO 209 and A236 encoding secreted or transmembrane  
 CC proteins. The polypeptides, nucleic acids and their modulators may be  
 CC useful for treating or modulating cholesterol uptake, blood coagulation,  
 CC to modulate cell proliferation, morphogenesis and fate specification,  
 CC tissue repair and renewal, to treat cancer and promote wound healing,  
 CC modulate angiogenesis, treat hypercholesterolemia, hemophilia, Marfan  
 CC syndrome, protein S deficiency, modulate allergic or inflammatory

CC response, acid secretion, tropic effects on gastrointestinal mucosa, and  
 CC promote ulcer healing, treat bone cancer, achondroplasia, myeloma,  
 CC fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma, osteoporosis,  
 CC leukemia, anemia, thalassemia, cerebral edema, hydrocephalus, brain  
 CC herniations, meningitis, ischemic brain or heart disease, infarction,  
 CC intracranial hemorrhage, pancreatitis, diabetes, angina, hypotensive heart  
 CC disease, pulmonary heart disease, rheumatic fever, congenital heart  
 CC disease, myocardial disease, atherosclerosis, hypertension, jaundice,  
 CC hepatic failure, cirrhosis, glomerulonephritis, Goodpasture's syndrome,  
 CC sickle cell disease, renal failure, ischemic bowel disease, Crohn's  
 CC disease, hernias, hypoadrenalism, hyperadrenalism, Cushing's syndrome,  
 CC neoplasia, pulmonary disorders, asthma, ovarian disorders, McCune  
 CC Albright syndrome, infertility, uterine disorders, viral disease. The  
 CC present sequence represents a consensus thyroglobulin type 1 domain  
 XX  
 SQ Sequence 51 AA;

Query Match 60.3%; Score 38; DB 4; Length 51;  
 Best Local Similarity 50.0%; Pred. No. 25;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFAP 12  
 : : ||||| :  
 Db 2 VPNCDENGFYKP 13

## RESULT 29

ID ABU47981 standard; protein; 197 AA.  
 XX

AC ABU47981;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #33508.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salmomella typhi.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.

XX N-PSDB; ACA51851.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 75905; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 197 AA;

Query Match 60.3%; Score 38; DB 6; Length 197;  
 Best Local Similarity 87.5%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PVLDEGL 9  
 : : ||||| :  
 Db 3 PVFDEGL 10

## RESULT 30

ID ADC01060 standard; protein; 197 AA.  
 XX

AC ADC01060;

XX 04-DEC-2003 (first entry)

XX Enterohaemorrhagic *E. coli* O157:H7-specific protein SEQ ID NO: 1104.

XX enterohaemorrhagic; anti-bacterial.

XX *Escherichia coli*; O157:H7.

XX JP2002355074-A.

XX 10-DEC-2002.

XX 24-JAN-2002; 2002JP-00015959.

XX 24-JAN-2001; 2001JP-00112010.

XX (UVTS-) UNIV TSUKUBA.

XX WPI; 2003-451640/43.

XX Enterohaemorrhagic *Escherichia coli* O157:H7-specific nucleic acid molecule  
 PT and a polypeptide and its use, a polypeptide, a vector and a host cell.  
 XX Claim 3; SEQ ID NO 1104; 2067pp; Japanese.

XX The invention relates to a novel enterohaemorrhagic *Escherichia coli*  
 CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention  
 CC has anti-bacterial activity. The polypeptide can be used in detection  
 CC and/or treatment of O157:H7 infection. The nucleotide sequence of the  
 CC genome of Enterohaemorrhagic *E. coli* O157:H7 was determined. The present  
 CC sequence represents an *E. coli* O157:H7-specific polypeptide of the

CC invention.  
XX SQ Sequence 197 AA;  
Query Match 60.3%; Score 38; DB 7; Length 197;  
Best Local Similarity 87.5%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 2 PVLDENGL 9  
Db 3 PVFDENGL 10  
RESULT 31  
ABP66209  
ID ABP66209 standard; protein; 213 AA.  
XX AC ABP66209;  
XX DT 19-NOV-2002 (first entry)  
XX DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:953.  
XX KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;  
XX KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;  
XX KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;  
XX KW rotavirus; food composition; pharmaceutical composition.  
XX OS Bifidobacterium longum.  
XX PN EPI227152-A1.  
XX PD 31-JUL-2002.  
XX PF 30-JAN-2001; 2001EP-00102050.  
XX PR 30-JAN-2001; 2001EP-00102050.  
XX PA (NEST ) SOC PROD NESTLE SA.  
XX DR WPI; 2002-668397/72.  
XX PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as  
PT a probe or primer for detecting and/or identifying Bifidobacterium longum  
PT in a biological sample.  
XX PS Claim 3; SEQ ID NO 953; 80pp; English.  
XX CC The present invention describes a polynucleotide (I) comprising a  
CC sequence of a Bifidobacterium genome selected from the nucleotide  
CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at  
CC least 90% identity or which hybridises with the sequences given in  
CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a  
CC fusion protein, comprising a sequence selected from 1097 sequences given  
CC in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a  
CC heterologous polypeptide. (II) has antidiarrheic and antibacterial  
CC activities, and can be used as an inhibitor of Salmonella. (I) (which is  
CC a probe) is useful for the detection and/or identification of  
CC Bifidobacterium longum in a biological sample. A carrier containing the  
CC lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM I-2618) can be  
CC used for preventing and/or treating diarrhoea brought about by pathogenic  
CC bacteria and/or rotavirus. The carrier is a food composition selected  
CC from milk, yogurt, curd, cheese, fermented milks, milk based  
CC products, ice-creams, fermented cereal based products, milk based  
CC powders, infant formula, pet food or a pharmaceutical composition  
CC selected from tablets, liquid bacterial suspensions, dried oral  
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.  
CC (I) is useful in DNA arrays or chips to carry out analysis of the  
CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent  
CC Bifidobacterium related nucleotide sequences given in the Sequence  
CC Listing from the present invention but not mentioned further within the  
CC specification. N.B. The sequence data for this patent is not represented  
CC in the printed specification but is based on sequence information

CC supplied by the European Patent Office  
XX SQ Sequence 213 AA;  
Query Match 60.3%; Score 38; DB 5; Length 213;  
Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 3 VLDENGLFAP 12  
Db 17 VWDENGFVP 26  
RESULT 32  
AAG82332  
ID AAG82332 standard; protein; 313 AA.  
XX AC AAG82332;  
XX DT 03-SEP-2001 (first entry)  
XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:1758.  
XX KW Staphylococcus epidermidis S1 strain; infection; diagnosis; vaccination;  
XX KW endocarditis.  
XX OS Staphylococcus epidermidis.  
XX PN WO200134809-A2.  
XX PD 17-MAY-2001.  
XX PF 09-NOV-2000; 2000WO-US030782.  
XX PR 09-NOV-1999; 99US-0164258P.  
XX PA (GLAX ) GLAXO GROUP LTD.  
XX PI Kimmerly WJ;  
XX DR WPI; 2001-316495/33.  
XX DR N-PSDB; AAH53182.  
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
PT useful for vaccinating against infections, e.g. endocarditis.  
XX PS Claim 18; Page 486; 2188pp; English.  
XX CC AAH53304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)  
CC and (II) can have antibacterial activity and therefore can be used in  
CC vaccination. The nucleic acids (II) may be used to produce the S.  
CC epidermidis polypeptides (II) via the production of vectors containing  
CC them which are used to produce hosts cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the  
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA  
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098  
CC represent oligonucleotide sequences and primers which are used in the  
CC exemplification of the present invention. N.B. The present invention  
CC specifically claims all the polynucleotide sequences given in the  
CC sequence listing of the present specification, however the sequence  
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
CC for SEQ ID NO:4455 to 4464  
XX SQ Sequence 313 AA;  
Query Match 60.3%; Score 38; DB 4; Length 313;  
Best Local Similarity 77.8%; Pred. No. 2.1e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PVLDELGLF 10  
 Db 64 PILDELGLF 72

RESULT 33  
 ABR52942  
 ID ABR52942 standard; protein; 322 AA.  
 XX  
 AC ABR52942;  
 XX  
 DT 20-JUN-2003 (first entry)  
 XX  
 DE Protein sequence #SEQ ID 749.  
 XX  
 KW Multiprotein complex; eukaryote; drug target; diagnosis.  
 XX  
 OS Saccharomyces cerevisiae.  
 PN EPI258494-A1.  
 XX  
 PD 20-NOV-2002.  
 XX  
 PF 20-DEC-2001; 2001EP-00130253.  
 XX  
 PR 15-MAY-2001; 2001EP-00111774.  
 XX  
 PA (CELL-) CELLZONE AG.  
 XX  
 PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;  
 PI Marzioch M, Schultz JD, Superti-Furga GD;  
 XX  
 DR WPI; 2003-250078/25.  
 DR N-PSDB; ACC60984.  
 XX  
 PT New isolated protein complexes useful for diagnosing a disease or  
 disorder, or as a target for an active agent of a pharmaceutical,  
 PT preferably a drug target in the treatment or prevention of disease or  
 disorder.  
 XX  
 PS Disclosure; SEQ ID NO 749; 17pp + Sequence Listing; English.  
 XX  
 CC The invention relates to multiprotein complexes from eukaryotes. Proteins  
 of the invention and DNA sequences encoding them are given in records  
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are  
 CC obtainable by using a protein as a bait and isolating the set of proteins  
 CC which is attached thereto from cells. Such protein complexes may comprise  
 CC up to 30 distinct proteins. Protein complexes of the invention are useful  
 CC for diagnosing a disease or disorder, or as a target for an active agent  
 CC of a pharmaceutical, preferably a drug target in the treatment or  
 CC prevention of a disease or disorder. Note: The sequence data for this  
 CC patent is not represented in the printed specification, but is based on  
 CC sequence information supplied by the European Patent Office. The complete  
 CC document is available on CD-ROM  
 XX  
 SQ Sequence 322 AA;

Query Match 60.3%; Score 38; DB 6; Length 322;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENG 8  
 Db 223 VPIIDENG 230

RESULT 34  
 ADK62016  
 ID ADK62016 standard; protein; 322 AA.  
 XX  
 AC ADK62016;

Query Match 60.3%; Score 38; DB 7; Length 322;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENG 8  
 Db 223 VPIIDENG 230

RESULT 35  
 ABP38384  
 ID ABP38384 standard; protein; 328 AA.  
 XX  
 AC ABP38384;  
 XX  
 DT 24-JUL-2002 (first entry)  
 XX  
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3229.  
 XX

XX 06-MAY-2004 (first entry)  
 DT  
 XX Disease treating protein complex-derived protein #134.  
 DE  
 XX protein complex; drug target; diagnosis.  
 KW  
 XX Unidentified.  
 OS  
 XX EPI338608-A2.  
 PN  
 XX 27-AUG-2003.  
 PD  
 XX 20-DEC-2002; 2002EP-00102502.  
 PF  
 XX 20-DEC-2001; 2001EP-00130253.  
 PR  
 XX (CELL-) CELLZONE AG.  
 PA  
 XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;  
 PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;  
 PI Michon A, Leutwein C, Rick J;  
 XX  
 DR WPI; 2003-638460/61.  
 DR N-PSDB; ADK62017.  
 XX  
 CC New proteins and protein complexes from eukaryotes, useful as targets in  
 drug screening, or in diagnosing or screening for the presence of a  
 CC disease or disorder, or a predisposition for developing a disease or  
 CC disorder in a subject.  
 XX  
 PS Disclosure; SEQ ID NO 267; 13pp; English.  
 XX  
 CC The invention relates to novel protein complexes comprising a first and a  
 second protein, or its derivative, fragment, homologue or variant. The  
 CC proteins are selected from given protein complexes, which are not defined  
 CC in the specification. The variants are encoded by nucleic acids that  
 CC hybridize to the nucleic acids encoding the proteins under low stringency  
 CC conditions. The protein complexes are useful as targets for an active  
 CC agent of a pharmaceutical. These protein complexes are particularly  
 CC useful as drugs targets for the treatment or preventing of a disease or  
 CC disorder. The complexes and methods above are useful in diagnosing or  
 CC screening for the presence of a disease or disorder or a predisposition  
 CC for developing a disease or disorder in a subject. These are also useful  
 CC in screening for a drug for treatment or prevention of a disease or  
 CC disorder. The molecule that modulates the amount, activity or protein  
 CC components of the complex is useful for the manufacture of a medicament  
 CC for the treatment or prevention of a disease or disorder. This sequence  
 CC corresponds to a protein of the invention. (Note: the sequence data for  
 CC this patent did not form part of the printed specification but was  
 CC obtained from the EPO in electronic format).  
 XX  
 SQ Sequence 322 AA;

KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 XX antibacterial; gene therapy.  
 OS Staphylococcus epidermidis.  
 XX US6380370-B1.  
 XX 30-APR-2002.  
 XX 13-AUG-1998; 98US-00134001.  
 XX 14-AUG-1997; 97US-0055779P.  
 PR 08-NOV-1997; 97US-0064964P.  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 XX Doucette-Stamm LA, Bush D;  
 XX WPI; 2002-381255/41.  
 DR N-PSDB; ABN90929.  
 XX Novel isolated nucleic acid encoding a Staphylococcus epidermis  
 PT polypeptide, useful for diagnosing and treating bacterial infections.  
 XX Disclosure; SEQ ID NO 3229; 267pp; English.  
 CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
 CC antibacterial activity and can be used in gene therapy. The sequences can  
 CC also be used in the diagnosis and treatment of bacterial infections,  
 CC particularly S. epidermidis infections. The sequences can be used to  
 CC screen for compounds able to interfere with the S. epidermidis life cycle  
 CC or inhibit S. epidermidis infection. N.B. The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the USPTO web site  
 XX Sequence 328 AA;  
 SQ  
 Query Match 60.3%; Score 38; DB 5; Length 328;  
 Best Local Similarity 77.8%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PVLDENGLF 10  
 Db :|||:|  
 79 PILDGLGLF 87  
 RESULT 36  
 ABU20549  
 ID ABU20549 standard; protein; 483 AA.  
 XX AC  
 XX ABU20549;  
 XX 19-JUN-2003 (first entry)  
 XX Protein encoded by Prokaryotic essential gene #6076.  
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX Bacteroides fragilis.  
 XX WO20027183-A2.  
 XX 03-OCT-2002.  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA24419.  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 25; SEQ ID NO 48473; 1766pp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC on a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than S. aureus, S. typhimurium,  
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 483 AA;  
 SQ  
 Query Match 60.3%; Score 38; DB 6; Length 483;  
 Best Local Similarity 58.3%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 IPVLDENGLFAP 12  
 Db :|||:|:|  
 343 VPVLKEHGVAP 354  
 RESULT 37  
 AAG31867  
 ID AAG31867 standard; protein; 515 AA.  
 XX AC  
 XX AAG31867;  
 XX 17-OCT-2000 (first entry)  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 38342.  
 DE Arabidopsis thaliana.  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 XX Arabidopsis thaliana.  
 OS

XX  
PN EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-01231180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139859P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140635P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151338P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.

```
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0158458P.
PR 29-SEP-1999; 99US-0158596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0158294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 60.3%; Score 38; DB 3; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVLDENG 8
Db 195 PVLDENG 201

RESULT 38
AAG31866
ID AAG31866 standard; protein; 582 AA.
XX
AC AAG31866;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38341.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 99US-0121825P.
XX 99US-0123180P.
XX 99US-0123548P.
XX 99US-0125788P.
XX 99US-0126284P.
XX 99US-0126785P.
XX 99US-0127462P.
XX 99US-0128234P.
XX 99US-0128714P.
XX 99US-0129845P.
XX 99US-0130077P.
XX 99US-0130449P.
XX 99US-0130510P.
XX 99US-0130891P.
XX 99US-0131449P.
XX 99US-0132048P.
XX 99US-0132407P.
XX 99US-0132484P.
XX 99US-0132485P.
XX 99US-0132486P.
XX 99US-0132487P.
XX 99US-0132863P.
XX 99US-0134218P.
XX 99US-0134219P.
XX 99US-0134370P.
XX 99US-0134768P.
XX 99US-0134941P.
XX 99US-0135124P.
XX 99US-0135353P.
XX 99US-0135629P.
XX 99US-0136021P.
XX 99US-0136392P.
XX 99US-0136782P.
XX 99US-0137222P.
XX 99US-0137528P.
XX 99US-0137502P.
XX 99US-0137724P.
XX 99US-0138094P.
XX 99US-0138540P.
XX 99US-0138847P.
XX 99US-0139119P.
XX 99US-0139452P.
XX 99US-0139453P.
XX 99US-0139492P.
XX 99US-0139454P.
XX 99US-0139455P.
XX 99US-0139456P.
XX 99US-0139457P.
XX 99US-0139458P.
XX 99US-0139459P.
XX 99US-0139460P.
XX 99US-0139461P.
XX 99US-0139462P.
XX 99US-0139463P.
XX 99US-0139763P.
XX 99US-0139750P.
XX 99US-0139817P.
XX 99US-0139899P.
XX 99US-0140353P.
XX 99US-0140354P.
XX 99US-0140695P.
XX 99US-0140823P.
XX 99US-0140991P.
XX 99US-0141287P.
XX 99US-0141842P.
XX 99US-0142154P.
XX 99US-0142055P.
XX 99US-0142390P.
XX 99US-0142803P.
XX 99US-0142920P.
XX 99US-0142977P.
```

PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 17-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144333P.  
PR 21-JUL-1999; 99US-0144334P.  
PR 22-JUL-1999; 99US-0144335P.  
PR 23-JUL-1999; 99US-0144336P.  
PR 24-JUL-1999; 99US-0144337P.  
PR 25-JUL-1999; 99US-0144338P.  
PR 26-JUL-1999; 99US-0144339P.  
PR 27-JUL-1999; 99US-0144340P.  
PR 28-JUL-1999; 99US-0144341P.  
PR 29-JUL-1999; 99US-0144342P.  
PR 30-JUL-1999; 99US-0144343P.  
PR 31-JUL-1999; 99US-0144344P.  
PR 01-AUG-1999; 99US-0144345P.  
PR 02-AUG-1999; 99US-0144346P.  
PR 03-AUG-1999; 99US-0144347P.  
PR 04-AUG-1999; 99US-0144348P.  
PR 05-AUG-1999; 99US-0144349P.  
PR 06-AUG-1999; 99US-0144350P.  
PR 07-AUG-1999; 99US-0144351P.  
PR 08-AUG-1999; 99US-0144352P.  
PR 09-AUG-1999; 99US-0144353P.  
PR 10-AUG-1999; 99US-0144354P.  
PR 11-AUG-1999; 99US-0144355P.  
PR 12-AUG-1999; 99US-0144356P.  
PR 13-AUG-1999; 99US-0144357P.  
PR 14-AUG-1999; 99US-0144358P.  
PR 15-AUG-1999; 99US-0144359P.  
PR 16-AUG-1999; 99US-0144360P.  
PR 17-AUG-1999; 99US-0144361P.  
PR 18-AUG-1999; 99US-0144362P.  
PR 19-AUG-1999; 99US-0144363P.  
PR 20-AUG-1999; 99US-0144364P.  
PR 21-AUG-1999; 99US-0144365P.  
PR 22-AUG-1999; 99US-0144366P.  
PR 23-AUG-1999; 99US-0144367P.  
PR 24-AUG-1999; 99US-0144368P.  
PR 25-AUG-1999; 99US-0144369P.  
PR 26-AUG-1999; 99US-0144370P.  
PR 27-AUG-1999; 99US-0144371P.  
PR 28-AUG-1999; 99US-0144372P.  
PR 29-AUG-1999; 99US-0144373P.  
PR 30-AUG-1999; 99US-0144374P.  
PR 31-AUG-1999; 99US-0144375P.  
PR 01-SEP-1999; 99US-0144376P.  
PR 02-SEP-1999; 99US-0144377P.  
PR 03-SEP-1999; 99US-0144378P.  
PR 04-SEP-1999; 99US-0144379P.  
PR 05-SEP-1999; 99US-0144380P.  
PR 06-SEP-1999; 99US-0144381P.  
PR 07-SEP-1999; 99US-0144382P.  
PR 08-SEP-1999; 99US-0144383P.  
PR 09-SEP-1999; 99US-0144384P.  
PR 10-SEP-1999; 99US-0144385P.  
PR 11-SEP-1999; 99US-0144386P.  
PR 12-SEP-1999; 99US-0144387P.  
PR 13-SEP-1999; 99US-0144388P.  
PR 14-SEP-1999; 99US-0144389P.  
PR 15-SEP-1999; 99US-0144390P.  
PR 16-SEP-1999; 99US-0144391P.  
PR 17-SEP-1999; 99US-0144392P.  
PR 18-SEP-1999; 99US-0144393P.  
PR 19-SEP-1999; 99US-0144394P.  
PR 20-SEP-1999; 99US-0144395P.  
PR 21-SEP-1999; 99US-0144396P.  
PR 22-SEP-1999; 99US-0144397P.  
PR 23-SEP-1999; 99US-0144398P.  
PR 24-SEP-1999; 99US-0144399P.  
PR 25-SEP-1999; 99US-0144400P.  
PR 26-SEP-1999; 99US-0144401P.  
PR 27-SEP-1999; 99US-0144402P.  
PR 28-SEP-1999; 99US-0144403P.  
PR 29-SEP-1999; 99US-0144404P.

PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157118P.  
PR 06-OCT-1999; 99US-0157119P.  
PR 07-OCT-1999; 99US-0157120P.  
PR 08-OCT-1999; 99US-0157121P.  
PR 09-OCT-1999; 99US-0157122P.  
PR 10-OCT-1999; 99US-0157123P.  
PR 11-OCT-1999; 99US-0157124P.  
PR 12-OCT-1999; 99US-0157125P.  
PR 13-OCT-1999; 99US-0157126P.  
PR 14-OCT-1999; 99US-0157127P.  
PR 15-OCT-1999; 99US-0157128P.  
PR 16-OCT-1999; 99US-0157129P.  
PR 17-OCT-1999; 99US-0157130P.  
PR 18-OCT-1999; 99US-0157131P.  
PR 19-OCT-1999; 99US-0157132P.  
PR 20-OCT-1999; 99US-0157133P.  
PR 21-OCT-1999; 99US-0157134P.  
PR 22-OCT-1999; 99US-0157135P.  
PR 23-OCT-1999; 99US-0157136P.  
PR 24-OCT-1999; 99US-0157137P.  
PR 25-OCT-1999; 99US-0157138P.  
PR 26-OCT-1999; 99US-0157139P.  
PR 27-OCT-1999; 99US-0157140P.  
PR 28-OCT-1999; 99US-0157141P.  
PR 29-OCT-1999; 99US-0157142P.

Query Match 60.3%; Score 38; DB 3; Length 582;

Best Local Similarity 100.0%; Pred. No. 4.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PVLDENG 8

Db 262 PVLDENG 268

RESULT 39

ADD27928

ID ADD27928 standard; protein; 613 AA.

XX AC ADD27928;

XX DT 15-JAN-2004 (first entry)

XX DE Human hCOX-3 (cs) reading frame 1 protein SEQ ID NO:14.

XX KW cyclooxxygenase type 1; cyclooxxygenase type 1 variant protein;

XX KW COX-1 variant protein; genetic disease; tissue typing;

XX OS forensic identification; human; COX-3.

XX OS Homo sapiens.

XX PN W02003029411-A2.

XX PD 10-APR-2003.

XX PF 28-SEP-2002; 2002WO-US030947.

XX PR 28-SEP-2001; 2001US-0326133P.

XX PR 15-APR-2002; 2002US-0373225P.

XX PR 16-APR-2002; 2002US-0373661P.

XX PR 16-SEP-2002; 2002US-0411575P.

XX PA (UYYO ) UNIV BRIGHAM YOUNG.

XX



PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147152P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150804P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157855P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161921P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 60.3%; Score 38; DB 3; Length 620;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PVLDENG 8  
Db 300 PVLDENG 306  
|||||

Search completed: January 7, 2005, 10:05:30  
Job time : 91 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2005, 10:02:01 ; Search time 26 Seconds  
(without alignments)  
30.608 Million cell updates/sec

Title: US-09-699-224A-1  
Perfect score: 63  
Sequence: 1 IPVLDEGLFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	41	65.1	64	4	US-09-328-352-6518	Sequence 6518, Ap
2	40	63.5	441	4	US-09-328-352-7754	Sequence 7754, Ap
3	38	60.3	313	4	US-09-710-279-1758	Sequence 1758, Ap
4	38	60.3	322	3	US-09-359-161-7	Sequence 7, Appl
5	38	60.3	322	4	US-09-538-092-297	Sequence 297, Appl
6	38	60.3	328	3	US-09-134-001C-3229	Sequence 3229, Ap
7	38	60.3	588	4	US-09-248-796A-20295	Sequence 20295, A
8	38	60.3	712	4	US-09-708-426-9	Sequence 9, Appl
9	37.5	59.5	1646	4	US-09-583-110-4667	Sequence 4667, Ap
10	37	58.7	104	3	US-08-946-329A-97	Sequence 97, Appl
11	37	58.7	174	4	US-09-903-456-40	Sequence 40, Appl
12	37	58.7	178	4	US-09-145-828A-24	Sequence 24, Appl
13	37	58.7	178	4	US-09-903-456-26	Sequence 26, Appl
14	37	58.7	219	4	US-09-903-456-38	Sequence 38, Appl
15	37	58.7	221	4	US-09-543-681A-4514	Sequence 4514, Ap
16	37	58.7	241	4	US-09-248-796A-15234	Sequence 15234, A
17	37	58.7	280	4	US-09-145-828A-21	Sequence 21, Appl
18	37	58.7	280	4	US-09-903-456-28	Sequence 28, Appl
19	37	58.7	286	4	US-09-903-456-59	Sequence 59, Appl
20	37	58.7	289	4	US-09-145-828A-17	Sequence 17, Appl
21	37	58.7	289	4	US-09-903-456-21	Sequence 21, Appl
22	37	58.7	289	4	US-09-903-456-34	Sequence 34, Appl
23	37	58.7	291	4	US-09-903-456-36	Sequence 36, Appl
24	37	58.7	293	4	US-09-145-828A-12	Sequence 12, Appl
25	37	58.7	293	4	US-09-903-456-19	Sequence 19, Appl
26	37	58.7	301	4	US-09-903-456-33	Sequence 33, Appl
27	37	58.7	317	4	US-09-145-828A-7	Sequence 7, Appl

28	37	58.7	317	4	US-09-903-456-13	Sequence 13, Appl
29	37	58.7	318	4	US-09-145-828A-19	Sequence 19, Appl
30	37	58.7	318	4	US-09-903-456-25	Sequence 25, Appl
31	37	58.7	367	4	US-09-489-039A-10054	Sequence 10054, A
32	37	58.7	409	4	US-09-248-796A-18059	Sequence 18059, A
33	37	58.7	436	4	US-10-140-002-404	Sequence 404, App
34	36	57.1	209	4	US-09-540-236-3512	Sequence 3512, Ap
35	36	57.1	301	4	US-09-270-767-46419	Sequence 46419, A
36	36	57.1	403	4	US-08-311-731A-104	Sequence 104, App
37	36	57.1	506	4	US-09-270-767-41642	Sequence 41642, A
38	36	57.1	566	4	US-09-543-681A-4544	Sequence 4544, Ap
39	36	57.1	593	4	US-09-252-991A-20441	Sequence 20441, A
40	36	57.1	1428	4	US-09-252-991A-30731	Sequence 30731, A
41	36	57.1	7257	3	US-09-335-409-5	Sequence 5, Appli
42	36	57.1	7257	3	US-09-568-102-5	Sequence 5, Appli
43	36	57.1	7257	3	US-09-567-969-5	Sequence 5, Appli
44	36	57.1	7257	3	US-09-568-480-5	Sequence 5, Appli
45	36	57.1	7257	3	US-09-568-486-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-09-328-352-6518  
; Sequence 6518, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6518  
; LENGTH: 64  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6518

Query Match 65.1%; Score 41; DB 4; Length 64;  
Best Local Similarity 80.0%; Pred. No. 1.8;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VLDEGLFAP 12  
| | | | | | |  
Db 36 VLDEGLLKP 45

RESULT 2  
US-09-328-352-7754  
; Sequence 7754, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7754  
; LENGTH: 441  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7754

Query Match 63.5%; Score 40; DB 4; Length 441;  
Best Local Similarity 58.3%; Pred. No. 27;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLFAP 12

Db 118 LPTYDQVGLFAP 129  
:| | | | |

## RESULT 3

US-09-710-279-1758  
; Sequence 1758, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1758  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-1758

Query Match 60.3%; Score 38; DB 4; Length 313;  
Best Local Similarity 77.8%; Pred. No. 43;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PVLDELGLF 10  
:| | | | |

Db 64 PILDELGLF 72

## RESULT 4

US-09-359-161-7  
; Sequence 7, Application US/09359161A  
; Patent No. 6342656  
; GENERAL INFORMATION:  
; APPLICANT: Bradford, Kent J.  
; APPLICANT: Dahal, Peetambar  
; APPLICANT: Yang, Hong  
; APPLICANT: Cooley, Michael  
; APPLICANT: Downie, Bruce  
; APPLICANT: Gee, Oliver  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses  
; TITLE OF INVENTION: to Stress Conditions in Plants  
; FILE REFERENCE: 023070-095900US  
; CURRENT APPLICATION NUMBER: US/09/359,161A  
; CURRENT FILING DATE: 1999-07-21  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 322  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
; FEATURE:  
; OTHER INFORMATION: yeast sucrose nonfermenting protein kinase 1  
; OTHER INFORMATION: kinase subunit (SNF1)  
US-09-359-161-7

Query Match 60.3%; Score 38; DB 3; Length 322;  
Best Local Similarity 62.5%; Pred. No. 44;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENG 8  
:| | | | |

Db 223 VPIIDENG 230

## RESULT 5

US-09-538-092-297  
; Sequence 297, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurapatSeqFormatter Version 0.9  
; SEQ ID NO 297  
; LENGTH: 322  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number YGL115W  
US-09-538-092-297

Query Match 60.3%; Score 38; DB 4; Length 322;  
Best Local Similarity 62.5%; Pred. No. 44;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENG 8  
:| | | | |

Db 223 VPIIDENG 230

## RESULT 6

US-09-134-001C-3229  
; Sequence 3229, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3229  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3229

Query Match 60.3%; Score 38; DB 3; Length 328;  
Best Local Similarity 77.8%; Pred. No. 45;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PVLDELGLF 10  
:| | | | |

Db 79 PILDELGLF 87

## RESULT 7

US-09-248-796A-20295  
; Sequence 20295, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
;; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: 107196.132  
;; CURRENT APPLICATION NUMBER: US/09/248,796A  
;; PRIOR FILING DATE: 1999-02-12  
;; PRIOR APPLICATION NUMBER: US 60/074,725  
;; PRIOR FILING DATE: 1998-02-13  
;; PRIOR APPLICATION NUMBER: US 60/096,409  
;; PRIOR FILING DATE: 1998-08-13  
;; NUMBER OF SEQ ID NOS: 28208  
;; SEQ ID NO 20295  
;; LENGTH: 588  
;; TYPE: PRT  
;; ORGANISM: Candida albicans  
US-09-248-796A-20295

Query Match 60.3%; Score 38; DB 4; Length 588;  
Best Local Similarity 70.0%; Pred. No. 90;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 10  
Db 525 IPVVDLSLF 534  
|||:|:|

RESULT 8  
US-09-708-426-9  
; Sequence 9, Application US/09708426  
; Patent No. 644429  
; GENERAL INFORMATION:  
; APPLICANT: HAN, YE-SUN  
; APPLICANT: YU, YEON-GYU  
; APPLICANT: LIM, JAE-HWAN  
; TITLE OF INVENTION: GENE CODING FOR DNA LIGASE OF HYPERTHERMOPHILIC BACTERIA AQUIFEX  
; FILE REFERENCE: 199579US0  
; CURRENT APPLICATION NUMBER: US/09/708,426  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: KR99-49591  
; PRIOR FILING DATE: 1999-11-10  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Rhodothermus marinus  
US-09-708-426-9

Query Match 60.3%; Score 38; DB 4; Length 712;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDEGL 9  
Db 697 IPMLDEGL 705  
|||:|:|

RESULT 9  
US-09-583-110-4667  
; Sequence 4667, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553

;; PRIOR FILING DATE: 1997-07-02  
;; NUMBER OF SEQ ID NOS: 5322  
;; SEQ ID NO 4667  
;; LENGTH: 1646  
;; TYPE: PRT  
;; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4667

Query Match 59.5%; Score 37.5; DB 4; Length 1646;  
Best Local Similarity 52.9%; Pred. No. 3.7e+02;  
Matches 9; Conservative 1; Mismatches 0; Indels 7; Gaps 1;

Qy 3 VLDEN-----GLFAP 12  
Db 444 ILDENGKRLSLGLFAP 460  
|||:|:|

RESULT 10  
US-08-946-329A-97  
; Sequence 97, Application US/08946329A  
; Patent No. 6057091  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; APPLICANT: Porter, Jeffrey A.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 109  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/946,329A  
; FILING DATE: 07-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/061,323  
; FILING DATE: 07-OCT-1996  
; APPLICATION NUMBER: 08/729,743  
; FILING DATE: 10-JUL-1996  
; APPLICATION NUMBER: 08/567,357  
; FILING DATE: 04-DEC-1995  
; APPLICATION NUMBER: 08/349,498  
; FILING DATE: 02-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/140001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 97:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 104 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-946-329A-97

Query Match 58.7%; Score 37; DB 3; Length 104;  
Best Local Similarity 72.7%; Pred. No. 18;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFA 11  
Db 40 IMMLDEGLVA 50  
|||:|:|

```
RESULT 11
US-09-903-456-40
; Sequence 26, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-40

Query Match      58.7%; Score 37; DB 4; Length 174;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10
Db 80 VPILARNGLF 89
      :||| |||

RESULT 12
US-09-145-828A-24
; Sequence 24, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.O1
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-145-828A-24

Query Match      58.7%; Score 37; DB 4; Length 178;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10
Db 3 VPILARNGLF 12
      :||| |||

RESULT 13
US-09-903-456-26
; Sequence 26, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-26

Query Match      58.7%; Score 37; DB 4; Length 178;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10
Db 10 VPILARNGLF 19
      :||| |||

RESULT 14
US-09-903-456-38
; Sequence 38, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-38

Query Match      58.7%; Score 37; DB 4; Length 219;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10
Db 10 VPILARNGLF 19
      :||| |||

RESULT 15
US-09-903-456-26
```

```
US-09-543-681A-4514
; Sequence 4514, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4514
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4514

Query Match      58.7%; Score 37; DB 4; Length 221;
Best Local Similarity 72.7%; Pred. No. 43;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDEGLFA 11
Db      63 IPVLDDQATLFA 73

RESULT 16
US-09-248-796A-15234
; Sequence 15234, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15234
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15234

Query Match      58.7%; Score 37; DB 4; Length 241;
Best Local Similarity 60.0%; Pred. No. 48;
Matches      6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      3 VLDEGLFAP 12
Db      214 LLDDQGLFSP 223

RESULT 17
US-09-145-828A-21
; Sequence 21, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.01
```

```
US-09-145-828A-21
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (280)...(280)
; OTHER INFORMATION: Xaa = Unknown or other at position 280
US-09-145-828A-21

Query Match      58.7%; Score 37; DB 4; Length 280;
Best Local Similarity 60.0%; Pred. No. 57;
Matches      6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDEGLF 10
Db      70 VPILARNGLF 79

RESULT 18
US-09-903-456-28
; Sequence 28, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (280)...(280)
; OTHER INFORMATION: Xaa = Unknown or Other at position 280
US-09-903-456-28

Query Match      58.7%; Score 37; DB 4; Length 280;
Best Local Similarity 60.0%; Pred. No. 57;
Matches      6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDEGLF 10
Db      70 VPILARNGLF 79

RESULT 19
US-09-903-456-59
; Sequence 59, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
```

; APPLICANT: Pereira, Suzette L.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P3  
; CURRENT APPLICATION NUMBER: US/09/903,456  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-09-903-456-59

Query Match 58.7%; Score 37; DB 4; Length 286;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 10

Db 77 VPILARNGLF 86

## RESULT 20

US-09-145-828A-17

; Sequence 17, Application US/09145828A

; Patent No. 6403349

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Mukerji, Pradip

; APPLICANT: Leonard, Amanda E. Y.

; APPLICANT: Huang, Yung-Sheng

; APPLICANT: Thurmond, Jennifer

; APPLICANT: Kirchner, Stephen J.

; APPLICANT: Parker-Barnes, Jennifer M.

; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF

; FILE REFERENCE: 6407.US.O1

; CURRENT APPLICATION NUMBER: US/09/145,828A

; CURRENT FILING DATE: 1998-09-02

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Mortierella alpina

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (289)...(289)

; OTHER INFORMATION: Xaa = Unknown or other at position 289

US-09-145-828A-17

Query Match 58.7%; Score 37; DB 4; Length 289;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 10

Db 79 VPILARNGLF 88

## RESULT 21

US-09-903-456-21

; Sequence 21, Application US/09903456

; Patent No. 6677145

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Mukerji, Pradip

; APPLICANT: Leonard, Amanda Eun-Yeong

; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Pereira, Suzette L.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P3  
; CURRENT APPLICATION NUMBER: US/09/903,456  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (289)...(289)  
; OTHER INFORMATION: Xaa = Unknown or Other at position 289  
US-09-903-456-21

Query Match 58.7%; Score 37; DB 4; Length 289;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 10

Db 79 VPILARNGLF 88

## RESULT 22

US-09-903-456-34

; Sequence 34, Application US/09903456

; Patent No. 6677145

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Mukerji, Pradip

; APPLICANT: Leonard, Amanda Eun-Yeong

; APPLICANT: Huang, Yung-Sheng

; APPLICANT: Pereira, Suzette L.

; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF

; FILE REFERENCE: 6407.US.P3

; CURRENT APPLICATION NUMBER: US/09/903,456

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: US 09/624,670

; PRIOR FILING DATE: 2000-07-24

; PRIOR APPLICATION NUMBER: US 09/379,095

; PRIOR FILING DATE: 1999-08-23

; PRIOR APPLICATION NUMBER: US 09/145,828

; PRIOR FILING DATE: 1998-09-02

; NUMBER OF SEQ ID NOS: 116

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 34

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Mortierella alpina

US-09-903-456-34

Query Match 58.7%; Score 37; DB 4; Length 289;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 10

Db 80 VPILARNGLF 89

## RESULT 23

US-09-903-456-36

; Sequence 36, Application US/09903456

```

; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR FILING DATE: 2000-07-24
; PRIOR FILING DATE: 1999-08-23
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-36

Query Match      58.7%; Score 37; DB 4; Length 291;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDENGLF 10
Db      82 VPILARNGLF 91

RESULT 24
US-09-145-828A-12
; Sequence 12, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407 US 01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (293)...(293)
; OTHER INFORMATION: Xaa = Unknown or other at position 293
US-09-145-828A-12

Query Match      58.7%; Score 37; DB 4; Length 293;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDENGLF 10
Db      83 VPILARNGLF 92

RESULT 25
US-09-903-456-19

```

```

; Sequence 19, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR FILING DATE: 2000-07-24
; PRIOR FILING DATE: 1999-08-23
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (293)...(293)
; OTHER INFORMATION: Xaa = Unknown or Other at position 293
US-09-903-456-19

Query Match      58.7%; Score 37; DB 4; Length 293;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDENGLF 10
Db      83 VPILARNGLF 92

RESULT 26
US-09-903-456-33
; Sequence 33, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR FILING DATE: 2000-07-24
; PRIOR FILING DATE: 1999-08-23
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (301)...(301)
; OTHER INFORMATION: Xaa = Unknown or Other at position 301
US-09-903-456-33

Query Match      58.7%; Score 37; DB 4; Length 301;
Best Local Similarity 60.0%; Pred. No. 62;

```

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10  
:|:| ||||  
Db 91 VPILARNGLF 100

RESULT 27  
US-09-145-828A-7  
; Sequence 7, Application US/09145828A  
; Patent No. 6403349  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda E. Y.  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Thurmond, Jennifer  
; APPLICANT: Kirchner, Stephen J.  
; APPLICANT: Parker-Barnes, Jennifer M.  
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF  
; FILE REFERENCE: 6407.US.01  
; CURRENT APPLICATION NUMBER: US/09/145,828A  
; CURRENT FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 317  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-09-145-828A-7

Query Match 58.7%; Score 37; DB 4; Length 317;  
Best Local Similarity 60.0%; Pred. No. 66;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10  
:|:| ||||  
Db 108 VPILARNGLF 117

RESULT 28  
US-09-903-456-13  
; Sequence 13, Application US/09903456  
; Patent No. 6677145  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Pereira, Suzette L.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P3  
; CURRENT APPLICATION NUMBER: US/09/903,456  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 317  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-09-903-456-13

Query Match 58.7%; Score 37; DB 4; Length 317;  
Best Local Similarity 60.0%; Pred. No. 66;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10

Db 108 VPILARNGLF 117  
:|:| ||||

RESULT 29  
US-09-145-828A-19  
; Sequence 19, Application US/09145828A  
; Patent No. 6403349  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda E. Y.  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Thurmond, Jennifer  
; APPLICANT: Kirchner, Stephen J.  
; APPLICANT: Parker-Barnes, Jennifer M.  
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF  
; FILE REFERENCE: 6407.US.01  
; CURRENT APPLICATION NUMBER: US/09/145,828A  
; CURRENT FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
; NAME/KEY: VARIANT  
; LOCATION: (318)...(318)  
; OTHER INFORMATION: Xaa = Unknown or other at position 318  
US-09-145-828A-19

Query Match 58.7%; Score 37; DB 4; Length 318;  
Best Local Similarity 60.0%; Pred. No. 66;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10  
:|:| ||||  
Db 108 VPILARNGLF 117

RESULT 30  
US-09-903-456-25  
; Sequence 25, Application US/09903456  
; Patent No. 6677145  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Pereira, Suzette L.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P3  
; CURRENT APPLICATION NUMBER: US/09/903,456  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
; NAME/KEY: VARIANT  
; LOCATION: (318)...(318)  
; OTHER INFORMATION: Xaa = Unknown or Other at position 318  
US-09-903-456-25

```
Query Match      58.7%; Score 37; DB 4; Length 318;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDEGLF 10
      :||| |||
Db     108 VPILARGLF 117

RESULT 31
US-09-489-039A-10054
; Sequence 10054, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10054
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10054

Query Match      58.7%; Score 37; DB 4; Length 367;
Best Local Similarity 60.0%; Pred. No. 79;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDEGLF 10
      :||| |||
Db     297 VFLLDNGKF 306

RESULT 32
US-09-248-796A-18059
; Sequence 18059, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18059
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18059

Query Match      58.7%; Score 37; DB 4; Length 409;
Best Local Similarity 77.8%; Pred. No. 89;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 LDENGLFAP 12
      ||||| |||
Db     229 LDEDGLIAP 237

RESULT 33
US-10-140-002-404
; Sequence 404, Application US/10140002
; Patent No. 6725730
```

```
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 404
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-002-404

Query Match      58.7%; Score 37; DB 4; Length 436;
Best Local Similarity 50.0%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IPVLDEGLFAP 12
      |||: |||: |
Db     338 IPLCDEGYYKP 349

RESULT 34
US-09-540-236-3512
; Sequence 3512, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR.
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3512
; LENGTH: 209
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-3512

Query Match      57.1%; Score 36; DB 4; Length 209;
Best Local Similarity 70.0%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDEGLF 10
      |||: |||
Db     96 IGVDENGYP 105

RESULT 35
US-09-270-767-46419
; Sequence 46419, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
```

```
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46419
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46419

Query Match          57.1%; Score 36; DB 4; Length 301;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      3 VLDENGLFAP 12
Db      112 IASENGLFVP 121

RESULT 36
US-08-311-731A-104
; Sequence 104, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-104

Query Match          57.1%; Score 36; DB 4; Length 403;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 IPVLDENGL 9
Db      192 VPVLDQNG 200
```

## RESULT 37

US-09-270-767-41642  
; Sequence 41642, Application US/09270767  
; Patent No. 6703491

## ; GENERAL INFORMATION:

; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 41642  
; LENGTH: 506  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-41642

Query Match 57.1%; Score 36; DB 4; Length 506;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PVLDENGLP 10

Db 110 PVYDKNGVF 118

## RESULT 38

US-09-543-681A-4544  
; Sequence 4544, Application US/09543681A  
; Patent No. 6605709

## ; GENERAL INFORMATION:

; APPLICANT: GARY BRETTON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 4544  
; LENGTH: 566  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-4544

Query Match 57.1%; Score 36; DB 4; Length 566;

Best Local Similarity 63.6%; Pred. No. 2e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLFA 11

Db 500 VPVLDAGGLHA 510

## RESULT 39

US-09-252-991A-20441  
; Sequence 20441, Application US/09252991A  
; Patent No. 6551795

## ; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20441  
; LENGTH: 593  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (289),(474),(511),(523),(557)  
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.  
US-09-252-991A-20441

Query Match 57.1%; Score 36; DB 4; Length 593;  
Best Local Similarity 54.5%; Pred. No. 2.1e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFA 11  
:| |::|||  
Db 491 VPAEDDDGLFA 501

RESULT 40  
US-09-252-991A-30731  
; Sequence 30731, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30731  
; LENGTH: 1428  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30731

Query Match 57.1%; Score 36; DB 4; Length 1428;  
Best Local Similarity 72.7%; Pred. No. 6e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFA 11  
|||:| |  
Db 1221 IPVLEEIGLVA 1231

Search completed: January 7, 2005, 10:06:32  
Job time : 27 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2005, 10:05:37 ; Search time 469 Seconds  
(without alignments)  
9.226 Million cell updates/sec

Title: US-09-699-224A-1

Perfect score: 63

Sequence: 1 IPVLDEGLFAP 12

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1603904 seqs, 360571292 residues

Total number of hits satisfying chosen parameters: 1603904

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCN\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCUTS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	65.1	288	15	US-10-424-599-220159 Sequence 220159,
2	41	65.1	691	15	US-10-282-122A-47568 Sequence 47568, A
3	41	65.1	691	15	US-10-282-122A-50917 Sequence 50917, A
4	40	63.5	200	17	US-10-425-115-339892 Sequence 339892,
5	39	61.9	280	16	US-10-437-963-114931 Sequence 114931,
6	39	61.9	315	9	US-09-738-626-5787 Sequence 5787, Ap
7	39	61.9	403	14	US-10-369-493-19334 Sequence 19334, A
8	39	61.9	688	15	US-10-282-122A-49243 Sequence 49243, A
9	39	61.9	994	16	US-10-437-963-157708 Sequence 157708,
10	39	61.9	2893	10	US-09-882-227-522 Sequence 522, App
11	39	61.9	2893	15	US-10-282-122A-58673 Sequence 58673, A
12	38	60.3	188	15	US-10-424-599-245562 Sequence 245562,
13	38	60.3	197	15	US-10-282-122A-75905 Sequence 75905, A

14	38	60.3	483	15	US-10-282-122A-48473 Sequence 48473, A
15	38	60.3	613	14	US-10-260-937-14 Sequence 14, Appl
16	38	60.3	707	14	US-10-156-761-9979 Sequence 9979, Ap
17	38	60.3	894	15	US-10-282-122A-77256 Sequence 77256, A
18	38	60.3	1464	16	US-10-437-963-145171 Sequence 145171,
19	37.5	59.5	1659	16	US-10-474-776-249 Sequence 249, App
20	37	58.7	12	14	US-10-014-340-737 Sequence 737, App
21	37	58.7	12	14	US-10-014-338-9 Sequence 9, Appli
22	37	58.7	104	15	US-10-460-594-97 Sequence 97, Appl
23	37	58.7	105	15	US-10-335-977-8958 Sequence 8958, Ap
24	37	58.7	136	15	US-10-425-114-39021 Sequence 39021, A
25	37	58.7	174	9	US-09-903-456-40 Sequence 40, Appl
26	37	58.7	174	14	US-10-156-911-40 Sequence 40, Appl
27	37	58.7	174	14	US-10-408-736-37 Sequence 37, Appl
28	37	58.7	174	15	US-10-220-120-393 Sequence 393, App
29	37	58.7	178	9	US-09-903-456-26 Sequence 26, Appl
30	37	58.7	178	14	US-10-156-911-26 Sequence 26, Appl
31	37	58.7	178	14	US-10-408-736-23 Sequence 23, Appl
32	37	58.7	180	16	US-10-767-701-33319 Sequence 33319, A
33	37	58.7	196	16	US-10-220-481-9 Sequence 9, Appli
34	37	58.7	219	9	US-09-903-456-38 Sequence 38, Appl
35	37	58.7	219	14	US-10-156-911-38 Sequence 38, Appl
36	37	58.7	219	14	US-10-408-736-35 Sequence 35, Appl
37	37	58.7	236	16	US-10-408-765A-1320 Sequence 1320, Ap
38	37	58.7	236	17	US-10-370-715B-146 Sequence 146, App
39	37	58.7	280	9	US-09-903-456-28 Sequence 28, Appl
40	37	58.7	280	14	US-10-156-911-28 Sequence 28, Appl
41	37	58.7	280	14	US-10-408-736-25 Sequence 25, Appl
42	37	58.7	286	9	US-09-903-456-59 Sequence 59, Appl
43	37	58.7	286	14	US-10-156-911-59 Sequence 59, Appl
44	37	58.7	286	14	US-10-408-736-57 Sequence 57, Appl
45	37	58.7	289	9	US-09-903-456-21 Sequence 21, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-424-599-220159  
; Sequence 220159, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 220159  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_40832C.1.pep  
US-10-424-599-220159

Query Match 65.1%; Score 41; DB 15; Length 288;  
Best Local Similarity 77.8%; Pred. No. 68;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LDENGLFAP 12

Db 198 LDENGLFKP 206  
|||||

##### RESULT 2

US-10-282-122A-47568  
; Sequence 47568, Application US/10282122A  
; Publication No. US20040029129A1

```
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47568
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-47568

Query Match      65.1%; Score 41; DB 15; Length 691;
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1  IPVLDEGL 9
Db      674  IPVLDEGL 682

RESULT 3
US-10-282-122A-50917
; Sequence 50917, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50917
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-50917

Query Match      65.1%; Score 41; DB 15; Length 696;
Best Local Similarity 72.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1  IPVLDEGLFA 11
Db      682  VPVLDEGLKA 692

RESULT 4
US-10-425-115-339892
; Sequence 339892, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 339892
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_7314C.1.pep
US-10-425-115-339892

Query Match      63.5%; Score 40; DB 17; Length 200;
Best Local Similarity 58.3%; Pred. No. 68;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1  IPVLDEGLFAP 12
Db      65  VPVQTEHGLFVP 76
```

```
RESULT 5
US-10-437-963-114931
; Sequence 114931, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 114931
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(280)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18575C.1.pep
US-10-437-963-114931

Query Match      61.9%; Score 39; DB 16; Length 280;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10
   :||| |||
Db 90 VPVLTRNGLF 99

RESULT 6
US-09-738-626-5787
; Sequence 5787, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5787
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5787

Query Match      61.9%; Score 39; DB 9; Length 316;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 LDENGLEFAP 12
   :||| |||
Db 167 VDENGKEAP 175

RESULT 7
US-10-369-493-19334
; Sequence 19334, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19334
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-10-369-493-19334

Query Match      61.9%; Score 39; DB 14; Length 403;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENGL 9
   :||: |||
Db 192 VPVVDQNGL 200

RESULT 8
US-10-282-122A-49243
; Sequence 49243, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
```

```
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 49243
/ LENGTH: 688
/ TYPE: PRT
/ ORGANISM: Burkholderia fungorum
US-10-282-122A-49243

Query Match      61.9%; Score 39; DB 15; Length 688;
Best Local Similarity 77.8%; Pred. No. 4.1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENGL 9
Db 672 IPVLEDDGM 680

RESULT 9
US-10-437-963-157708
/ Sequence 157708, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 157708
/ LENGTH: 994
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_57252C.1.pep
US-10-437-963-157708

Query Match      61.9%; Score 39; DB 16; Length 994;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VLDENGLFAP 12
Db 395 VLMENGLFGP 404

RESULT 10
US-09-882-227-522
/ Sequence 522, Application US/09882227
/ Publication No. US20030158396A1
/ GENERAL INFORMATION:
/ APPLICANT: Kleanthous, Harold
/ APPLICANT: Al-Garawi, Amal
/ APPLICANT: Miller, Charles
/ APPLICANT: Tomb, Jean-Francois
```

```
/ APPLICANT: Ooomen, Raymond P.
/ TITLE OF INVENTION: Identification of Polynucleotides
/ TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the
/ TITLE OF INVENTION: Genome
/ FILE REFERENCE: 06132/047002
/ CURRENT APPLICATION NUMBER: US/09/882,227
/ CURRENT FILING DATE: 2001-06-15
/ PRIOR APPLICATION NUMBER: US 08/902,615
/ PRIOR FILING DATE: 1997-07-29
/ NUMBER OF SEQ ID NOS: 638
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 522
/ LENGTH: 2893
/ TYPE: PRT
/ ORGANISM: Helicobacter pylori
US-09-882-227-522

Query Match      61.9%; Score 39; DB 10; Length 2893;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLFAP 12
Db 2077 IPNLGKGLFAP 2089

RESULT 11
US-10-282-122A-58673
/ Sequence 58673, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 58673
/ LENGTH: 2893
/ TYPE: PRT
```

```
; ORGANISM: Helicobacter pylori
US-10-282-122A-58673

Query Match
Best Local Similarity 61.9%; Score 39; DB 15; Length 2893;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLFAP 12
||| : |||||
Db 2077 IPNLGKGLFAP 2088

RESULT 12
US-10-424-599-245562
; Sequence 245562, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245562
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63774C.1.pep
US-10-424-599-245562

Query Match
Best Local Similarity 60.3%; Score 38; DB 15; Length 188;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VLDENGLFAP 12
||||| : |||
Db 45 VLDENGLPVP 54

RESULT 13
US-10-282-122A-75905
; Sequence 75905, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75905
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-75905

Query Match
Best Local Similarity 60.3%; Score 38; DB 15; Length 197;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PVLDENGL 9
||| : |||
Db 3 PVFDENGL 10

RESULT 14
US-10-282-122A-48473
; Sequence 48473, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48473
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Bacteroides fragilis
US-10-282-122A-48473

Query Match      60.3%; Score 38; DB 15; Length 483;
Best Local Similarity 58.3%; Pred. No. 4.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDEGLFAP 12
   :|||:|:|:|
Db 343 VPVLKEHGVAP 354

RESULT 15
US-10-260-937-14
; Sequence 14, Application US/10260937
; Publication No. US20030220306A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Daniel
; APPLICANT: Chandrasekharan, N. Vishvanath
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 07913-007001
; CURRENT APPLICATION NUMBER: US/10/260,937
; CURRENT FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-260-937-14

Query Match      60.3%; Score 38; DB 14; Length 613;
Best Local Similarity 58.3%; Pred. No. 5.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDEGLFAP 12
   :|||:|:|:|
Db 84 LPVLPHGLFRP 95

RESULT 16
US-10-156-761-9979
; Sequence 9979, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77256
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77256

Query Match      60.3%; Score 38; DB 15; Length 894;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVLDENG 8
   :|||:|:|:|
Db 720 PVLDENG 726

; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9979
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9979

Query Match      60.3%; Score 38; DB 14; Length 707;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDEGLFAP 12
   :|||:|:|:|
Db 290 VPVLDDRGQMTTP 301

RESULT 17
US-10-282-122A-77256
; Sequence 77256, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITEPA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77256
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77256

Query Match      60.3%; Score 38; DB 15; Length 894;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVLDENG 8
   :|||:|:|:|
Db 720 PVLDENG 726
```

; TITLE OF INVENTION: Diagnosis and Treatment of Alzheimer's Disease  
; FILE REFERENCE: 9195-078  
; CURRENT APPLICATION NUMBER: US/10/014,340  
; CURRENT FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 737  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-014-340-737

Query Match 58.7%; Score 37; DB 14; Length 12;  
Best Local Similarity 87.5%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDENG 8  
||| ||||  
Db 3 IPTVDENG 10

RESULT 21  
US-10-014-338-9  
; Sequence 9, Application US/10014338  
; Publication No. US20030092614A1  
; GENERAL INFORMATION:  
; APPLICANT: Herath, et al.  
; TITLE OF INVENTION: ADPI-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE AND  
; FILE REFERENCE: 9195-077  
; CURRENT APPLICATION NUMBER: US/10/014,338  
; CURRENT FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: 10/014,338  
; PRIOR FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-014-338-9

Query Match 58.7%; Score 37; DB 14; Length 12;  
Best Local Similarity 87.5%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDENG 8  
||| ||||  
Db 3 IPTVDENG 10

RESULT 22  
US-10-460-594-97  
; Sequence 97, Application US/10460594  
; Publication No. US20040018979A1  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; Porter, Jeffrey A.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 109  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:

RESULT 18  
US-10-437-963-145171  
; Sequence 145171, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 145171  
; LENGTH: 1464  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_45915C.1.pap  
US-10-437-963-145171

Query Match 60.3%; Score 38; DB 16; Length 1464;  
Best Local Similarity 87.5%; Pred. No. 1.4e+03;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 DENGFLAP 12  
||| |||  
Db 413 DENGFLAP 420

RESULT 19  
US-10-474-776-249  
; Sequence 249, Application US/10474776  
; Publication No. US20040110181A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLY  
; FILE REFERENCE: AM100649-PCT  
; CURRENT APPLICATION NUMBER: US/10/474,776  
; CURRENT FILING DATE: 2003-10-14  
; NUMBER OF SEQ ID NOS: 752  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 249  
; LENGTH: 1659  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-10-474-776-249

Query Match 59.5%; Score 37.5; DB 16; Length 1659;  
Best Local Similarity 52.9%; Pred. No. 2e+03;  
Matches 9; Conservative 1; Mismatches 0; Indels 7; Gaps 1;

Qy 3 VLDEN-----GLFAP 12  
: ||||  
Db 457 ILDENGKRLSLGLFAP 473

RESULT 20  
US-10-014-340-737  
; Sequence 737, Application US/10014340  
; Publication No. US20030064411A1  
; GENERAL INFORMATION:  
; APPLICANT: Herath, et al  
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and Uses Therefor, Including

```
/ APPLICATION NUMBER: US/10/460,594
/ FILING DATE: 11-Jun-2003
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/061,323
/ FILING DATE: 07-OCT-1997
/ APPLICATION NUMBER: 08/729,743
/ FILING DATE: 07-OCT-1996
/ APPLICATION NUMBER: 08/567,357
/ FILING DATE: 10-JUL-1996
/ APPLICATION NUMBER: 08/349,498
/ FILING DATE: 04-DEC-1995
/ APPLICATION NUMBER: <Unknown>
/ FILING DATE: 02-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Haile, Lisa A.
/ REGISTRATION NUMBER: 38,347
/ REFERENCE/DOCKET NUMBER: 07265/140001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619/678-5070
/ TELEFAX: 619/678-5099
/ INFORMATION FOR SEQ ID NO: 97:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 104 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-10-460-594-97

Query Match      58.7%; Score 37; DB 15; Length 104;
Best Local Similarity 72.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDEGLFA 11
      | : |||||
Db      40 IMWLDGLVA 50

RESULT 23
US-10-335-977-8958
/ Sequence 8958, Application US/10335977
/ Publication No. US20040052799A1
/ GENERAL INFORMATION:
/ APPLICANT: DOUGLAS SMITH et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
/ RELATING TO HELICOBACTER PYLORI FOR
/ DIAGNOSTICS AND THERAPEUTICS
/ NUMBER OF SEQUENCES: 10031
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 28 State Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-ROM ISO9660
/ COMPUTER: IBM PC Compatible
/ OPERATING SYSTEM: Windows NT 4.0
/ SOFTWARE: UNIX
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/335,977
/ FILING DATE: 30-Dec-2002
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/993,002
/ FILING DATE: 17-DEC-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mandragouras, Amy E.
/ REGISTRATION NUMBER: 36,207
/ REFERENCE/DOCKET NUMBER: GTN-018
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 227-7400
/ TELEFAX: (617) 742-4214
```

```
/ INFORMATION FOR SEQ ID NO: 8958:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 105 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Helicobacter pylori
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (B) LOCATION 1...105
/ SEQUENCE DESCRIPTION: SEQ ID NO: 8958:
US-10-335-977-8958

Query Match      58.7%; Score 37; DB 15; Length 105;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IPVLDEGLFAP 12
      | | | : | : |||
Db      14 IPNLGKGIFAP 25

RESULT 24
US-10-425-114-39021
/ Sequence 39021, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 39021
/ LENGTH: 136
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3158-010-D7_FLI.pap
US-10-425-114-39021

Query Match      58.7%; Score 37; DB 15; Length 136;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IPVLDEGLFAP 12
      | | | : | : |||
Db      89 IPVDDGGAKAP 100

RESULT 25
US-09-903-456-40
/ Sequence 40, Application US/09903456
/ Patent No. US20020138874A1
/ GENERAL INFORMATION:
/ APPLICANT: Abbott Laboratories
/ APPLICANT: Mukerji, Pradip
/ APPLICANT: Leonard, Amanda Eun-Yeong
/ APPLICANT: Huang, Yung-Sheng
/ APPLICANT: Pereira, Suzette L.
/ TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
/ FILE REFERENCE: 6407.US.P3
/ CURRENT APPLICATION NUMBER: US/09/903,456
/ CURRENT FILING DATE: 2001-07-11
/ PRIOR APPLICATION NUMBER: US 09/624,670
```

; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-09-903-456-40

Query Match 58.7%; Score 37; DB 9; Length 174;  
Best Local Similarity 60.0%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 2;

Qy 1 IPVLDEGLGF 10  
:|:| ||||  
Db 80 VPILARNGLF 89

## RESULT 26

US-10-156-911-40  
; Sequence 40, Application US/10156911  
; Publication No. US20030163845A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Pereira, Suzette L.

; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P4

; CURRENT APPLICATION NUMBER: US/10/156,911  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: US 09/903,456  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40

; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-10-156-911-40

Query Match 58.7%; Score 37; DB 14; Length 174;  
Best Local Similarity 60.0%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 2;

Qy 1 IPVLDEGLGF 10  
:|:| ||||  
Db 80 VPILARNGLF 89

## RESULT 27

US-10-408-736-37  
; Sequence 37, Application US/10408736  
; Publication No. US20030177508A1  
; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Das, Tapas  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Parker-Barnes, Jennifer M.  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Thurmond, Jennifer M.

; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P1  
; CURRENT APPLICATION NUMBER: US/10/408,736  
; CURRENT FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: US/09/379,095A  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-10-408-736-37

Query Match 58.7%; Score 37; DB 14; Length 174;  
Best Local Similarity 60.0%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 2;

Qy 1 IPVLDEGLGF 10  
:|:| ||||  
Db 80 VPILARNGLF 89

## RESULT 28

US-10-220-120-393  
; Sequence 393, Application US/10220120  
; Publication No. US20040048253A1  
; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: PANZER, Scott R.  
; APPLICANT: SPIRO, Peter A.  
; APPLICANT: BANVILLE, Steven C.  
; APPLICANT: SHAH, Purvi  
; APPLICANT: CHALUP, Michael S.  
; APPLICANT: CHANG, Simon C.  
; APPLICANT: CHEN, Alice  
; APPLICANT: D'SA, Steven A.  
; APPLICANT: AMSHEY, Stefan  
; APPLICANT: DAHL, Christopher R.  
; APPLICANT: DAM, Tam C.  
; APPLICANT: DANIELS, Susan E.  
; APPLICANT: DUFOUR, Gerard E.  
; APPLICANT: FLORES, Vincent  
; APPLICANT: FONG, Willy T.  
; APPLICANT: GREENAWALT, Lila B.  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: JONES, Anissa L.  
; APPLICANT: LIU, Tommy F.  
; APPLICANT: ROSEBERRY, Ann M.  
; APPLICANT: ROSEN, Bruce H.  
; APPLICANT: RUSSO, Frank D.  
; APPLICANT: STOCKREHER, Theresa K.  
; APPLICANT: DAFFO, Abel  
; APPLICANT: WRIGHT, Rachel J.  
; APPLICANT: YU, Jimmy Y.  
; APPLICANT: BRADLEY, Diana L.  
; APPLICANT: BRATCHER, Shawn R.  
; APPLICANT: CHEN, Wensheng  
; APPLICANT: COHEN, Howard J.  
; APPLICANT: HODGSON, David M.  
; APPLICANT: LINCOLN, Stephen E.  
; APPLICANT: JACKSON, Stuart

; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: PT-1113 PCT  
; CURRENT APPLICATION NUMBER: US/10/220,120  
; CURRENT FILING DATE: 2002-08-26  
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;  
; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;  
; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;  
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;

```
/ 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
/ 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
/ 60/205,324; 60/205,286
/ PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
/ 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
/ 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
/ 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
/ 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
/ 2000-05-16; 2000-05-17; 2000-05-16; 2000-05-16; 2000-05-17;
/ 2000-05-17; 2000-05-17
/ NUMBER OF SEQ ID NOS: 422
/ SOFTWARE: PERL Program
/ SEQ ID NO 393
/ LENGTH: 174
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: miac feature
/ OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:978427.5.orf2:2000FEB18
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 151
/ OTHER INFORMATION: unknown or other
US-10-220-120-393
```

```
Query Match 58.7%; Score 37; DB 15; Length 174;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 IPVLDENG 8
Db 104 IPVTDENG 111
```

```
RESULT 29
US-09-903-456-26
/ Sequence 26, Application US/09903456
/ Patent No. US20020138874A1
/ GENERAL INFORMATION:
/ APPLICANT: Abbott Laboratories
/ APPLICANT: Mukerji, Pradip
/ APPLICANT: Leonard, Amanda Eun-Yeong
/ APPLICANT: Huang, Yung-Sheng
/ APPLICANT: Pereira, Suzette L.
/ TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
/ FILE REFERENCE: 6407.US.P3
/ CURRENT APPLICATION NUMBER: US/09/903,456
/ PRIOR FILING DATE: 2001-07-11
/ PRIOR FILING DATE: 2000-07-24
/ PRIOR FILING DATE: 2000-07-24
/ PRIOR FILING DATE: 2000-07-24
/ PRIOR FILING DATE: 1999-08-23
/ PRIOR APPLICATION NUMBER: US 09/379,095
/ PRIOR APPLICATION NUMBER: US 09/145,828
/ PRIOR FILING DATE: 1998-09-02
/ NUMBER OF SEQ ID NOS: 116
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26
/ LENGTH: 178
/ TYPE: PRT
/ ORGANISM: Mortierella alpina
US-09-903-456-26
```

```
Query Match 58.7%; Score 37; DB 9; Length 178;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 IPVLDENGLF 10
Db 3 VPILARNGLF 12
```

```
RESULT 30
US-10-156-911-26
```

```
/ Sequence 26, Application US/10156911
/ Publication No. US20030163845A1
/ GENERAL INFORMATION:
/ APPLICANT: Abbott Laboratories
/ APPLICANT: Mukerji, Pradip
/ APPLICANT: Leonard, Amanda Eun-Yeong
/ APPLICANT: Huang, Yung-Sheng
/ APPLICANT: Pereira, Suzette L.
/ TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
/ FILE REFERENCE: 6407.US.P4
/ CURRENT APPLICATION NUMBER: US/10/156,911
/ CURRENT FILING DATE: 2002-10-01
/ PRIOR APPLICATION NUMBER: US 09/903,456
/ PRIOR FILING DATE: 2001-07-11
/ PRIOR APPLICATION NUMBER: US 09/624,670
/ PRIOR FILING DATE: 2000-07-24
/ PRIOR APPLICATION NUMBER: US 09/379,095
/ PRIOR FILING DATE: 1999-08-23
/ PRIOR APPLICATION NUMBER: US 09/145,828
/ PRIOR FILING DATE: 1998-09-02
/ NUMBER OF SEQ ID NOS: 122
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26
/ LENGTH: 178
/ TYPE: PRT
/ ORGANISM: Mortierella alpina
US-10-156-911-26
```

```
Query Match 58.7%; Score 37; DB 14; Length 178;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 IPVLDENGLF 10
Db 3 VPILARNGLF 12
```

```
RESULT 31
US-10-408-736-23
/ Sequence 23, Application US/10408736
/ Publication No. US20030177508A1
/ GENERAL INFORMATION:
/ APPLICANT: Abbott Laboratories
/ APPLICANT: Mukerji, Pradip
/ APPLICANT: Das, Tapas
/ APPLICANT: Huang, Yung-Sheng
/ APPLICANT: Parker-Barnes, Jennifer M.
/ APPLICANT: Leonard, Amanda Eun-yeong
/ APPLICANT: Thurmond, Jennifer M.
/ TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
/ FILE REFERENCE: 6407.US.P1
/ CURRENT APPLICATION NUMBER: US/10/408,736
/ CURRENT FILING DATE: 2003-04-04
/ PRIOR APPLICATION NUMBER: US/09/379,095A
/ PRIOR FILING DATE: 1999-08-23
/ PRIOR APPLICATION NUMBER: US 09/145,828
/ PRIOR FILING DATE: 1998-09-02
/ NUMBER OF SEQ ID NOS: 81
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 23
/ LENGTH: 178
/ TYPE: PRT
/ ORGANISM: Mortierella alpina
US-10-408-736-23
```

```
Query Match 58.7%; Score 37; DB 14; Length 178;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 IPVLDENGLF 10
Db 3 VPILARNGLF 12
```

```
RESULT 32
US-10-767-701-33319
; Sequence 33319, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 33319
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C29522_1.pep
US-10-767-701-33319

Query Match      58.7%; Score 37; DB 16; Length 180;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDENGLFP 12
Db      48 IPVLIQDGIFLP 59

RESULT 33
US-10-220-481-9
; Sequence 9, Application US/10220481
; Publication No. US20040110670A1
; GENERAL INFORMATION:
; APPLICANT: ARICO, Maria B., et al.
; TITLE OF INVENTION: Heterologous Expression of Neisserial Proteins
; FILE REFERENCE: CHIR-15883/01US
; CURRENT APPLICATION NUMBER: US/10/220,481
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: PCT/IB01/00452
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 633
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 9
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-220-481-9

Query Match      58.7%; Score 37; DB 16; Length 196;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDENGLGF 10
Db      136 VPLLDNNGKF 145

RESULT 34
US-09-903-456-38
; Sequence 38, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
```

```
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-903-456-38

Query Match      58.7%; Score 37; DB 9; Length 219;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDENGLGF 10
Db      10 VPILARNGLF 19

RESULT 35
US-10-156-911-38
; Sequence 38, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-10-156-911-38

Query Match      58.7%; Score 37; DB 14; Length 219;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IPVLDENGLGF 10
Db      10 VPILARNGLF 19

RESULT 36
US-10-408-736-35
; Sequence 35, Application US/10408736
; Publication No. US20030177508A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Das, Tapas
; APPLICANT: Huang, Yung-Sheng
```

; APPLICANT: Parker-Barnes, Jennifer M.  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Thurmond, Jennifer M.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407 US P1  
; CURRENT APPLICATION NUMBER: US/10/408,736  
; CURRENT FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: US/09/379,095A  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
US-10-408-736-35

Query Match 58.7%; Score 37; DB 14; Length 219;  
Best Local Similarity 60.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDEGLF 10  
Db 10 VPIARNGLF 19

RESULT 37  
US-10-408-765A-1320  
; Sequence 1320, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1320  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-1320

Query Match 58.7%; Score 37; DB 16; Length 236;  
Best Local Similarity 87.5%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDENG 8  
Db 119 IPVTDENG 126

RESULT 38  
US-10-370-715B-146  
; Sequence 146, Application US/10370715B  
; Publication No. US20040258678A1  
; GENERAL INFORMATION:  
; Patin Docket Preview  
; APPLICANT: BODARY, SARAH C.  
; APPLICANT: CLARK, HILLARY  
; APPLICANT: BRISDELL, HUNTE  
; APPLICANT: JACKMAN, JANET  
; APPLICANT: SCHOENFELD, JILL R.

; APPLICANT: WILLIAMS, P. MICKEY  
; APPLICANT: WOOD, WILLIAM I.  
; APPLICANT: WU, THOMAS D.  
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune  
; TITLE OF INVENTION: Related Diseases  
; FILE REFERENCE: P1948R1-US  
; CURRENT APPLICATION NUMBER: US/10/370,715B  
; CURRENT FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 742  
; SEQ ID NO 146  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-370-715B-146

Query Match 58.7%; Score 37; DB 17; Length 236;  
Best Local Similarity 87.5%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDENG 8  
Db 119 IPVTDENG 126

RESULT 39  
US-09-903-456-28  
; Sequence 28, Application US/09903456  
; Patent No. US20020138874A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda Eun-Yeong  
; APPLICANT: Huang, Yung-Sheng  
; APPLICANT: Pereira, Suzette L.  
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF  
; FILE REFERENCE: 6407.US.P3  
; CURRENT APPLICATION NUMBER: US/09/903,456  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US 09/624,670  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 09/379,095  
; PRIOR FILING DATE: 1999-08-23  
; PRIOR APPLICATION NUMBER: US 09/145,828  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Mortierella alpina  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (280)...(280)  
; OTHER INFORMATION: Xaa = Unknown or Other at position 280  
US-09-903-456-28

Query Match 58.7%; Score 37; DB 9; Length 280;  
Best Local Similarity 60.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDEGLF 10  
Db 70 VPIARNGLF 79

RESULT 40  
US-10-156-911-28  
; Sequence 28, Application US/10156911  
; Publication No. US20030163845A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Mukerji, Pradip  
; APPLICANT: Leonard, Amanda Eun-Yeong

```
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Mortierella alpina
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (280)...(280)
; OTHER INFORMATION: Xaa = Unknown or Other at position 280
US-10-156-911-28
```

```
Query Match          58.7%; Score 37; DB 14; Length 280;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 IPVLDENGLF 10
       :||| |||
Db      70 VPILARNGLF 79
```

```
Search completed: January 7, 2005, 10:19:19
Job time : 470 secs
```

Je Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2005, 10:02:01 ; Search time 25 Seconds  
(without alignments)  
46.184 Million cell updates/sec

Title: US-09-699-224A-1

Perfect score: 63

Sequence: 1 IPVLDENGLFP 12

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	71.4	359	2 C84983	hypothetical prote
2	41	65.1	180	2 A98321	hypothetical 15.0K
3	41	65.1	707	2 T35196	probable DNA gyras
4	40	63.5	157	1 F69185	yhcV homolog MTH64
5	40	63.5	337	2 AP2724	hypothetical prote
6	40	63.5	337	2 B97506	peptide ABC transp
7	39	61.9	140	2 H97350	two CBS domain con
8	39	61.9	450	2 AB3238	nitrotriacetate
9	39	61.9	700	2 G75318	DNA ligase - Delno
10	39	61.9	764	2 S64951	hypothetical prote
11	39	61.9	1621	2 A82255	hypothetical prote
12	39	61.9	2893	2 A64556	toxin-like outer m
13	38	60.3	162	2 T08505	trbH protein - Ent
14	38	60.3	197	2 AC0929	hypothetical prote
15	38	60.3	197	2 A90664	hypothetical prote
16	38	60.3	197	2 F85514	unknown protein fr
17	38	60.3	268	2 G69623	flagellar hook-bas
18	38	60.3	298	2 E97298	uncharacterized Fe
19	38	60.3	322	1 RGBYC3	regulatory protein
20	38	60.3	365	2 AF0864	membrane-bound lct
21	38	60.3	620	2 D84618	probable amino aci
22	38	60.3	894	2 E82211	DNA gyrase, chain
23	37.5	59.5	1659	2 G95057	endo-beta-N-acetyl
24	37.5	59.5	1659	2 H97926	hypothetical prote
25	37	58.7	144	2 F83988	inosine-5'-monopho
26	37	58.7	324	2 AB0982	probable 2-hydroxy
27	37	58.7	328	2 C65154	probable 2-hydroxy
28	37	58.7	328	2 F91183	probable dehydroge
29	37	58.7	328	2 B86030	probable dehydroge

30	37	58.7	347	2 AG2251	translation initia
31	37	58.7	357	2 I55210	tricarboxylate car
32	37	58.7	365	2 C85333	membrane-bound lct
33	37	58.7	365	2 A98088	membrane-bound lct
34	37	58.7	365	2 A65064	hypothetical prote
35	37	58.7	390	2 T44324	acyl-CoA thiolase
36	37	58.7	402	2 AC2638	probable acyl-CoA
37	37	58.7	402	2 B97420	serine/threonine-s
38	37	58.7	443	2 C85140	serine/threonine-s
39	37	58.7	443	2 S38327	hypothetical prote
40	37	58.7	468	2 T19628	hypothetical prote
41	37	58.7	504	2 A23282	RAD52 protein - ye
42	37	58.7	712	2 T48961	hypothetical prote
43	37	58.7	789	2 S33056	probable infected
44	37	58.7	809	2 B87260	sensory box/GGDEF
45	37	58.7	1051	2 C95367	conserved hypothet

ALIGNMENTS

RESULT 1

C84983  
hypothetical protein mlta [imported] - Buchnera sp. (strain APS)  
C;Species: Buchnera sp.  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C;Accession: C84983  
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
Nature 407, 81-86, 2000  
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A  
A;Reference number: A84930; MUID:20445173; PMID:10993077  
A;Accession: C84983  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-359 <STO>  
A;Cross-references: GB:AP000398; GSPDB:GN00144  
A;Experimental source: strain APS  
C;Genetics:  
A;Gene: mlta; BU458

Query Match 71.4%; Score 45; DB 2; Length 359;  
Best Local Similarity 70.0%; Pred. No. 2.1;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENGLF 10

Db 298 IPILDQGVF 307  
|||:|:|:|:  
|:|:|:|:|:

RESULT 2

A98321  
hypothetical 15.0K protein in cobo 3'region (orf6) [imported] - Agrobacterium tumefaciens  
C;Species: Agrobacterium tumefaciens  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: A98321  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: A98321  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-180 <KUR>  
A;Cross-references: UNIPROT:Q8UAS1; GB:AE007870; PIDN:AAK90091.1; PID:gi5160078; GSPDB:G

A;Gene: AGR\_L\_3039

A;Map position: linear chromosome

C;Superfamily: Transporter DME family

Query Match 65.1%; Score 41; DB 2; Length 180;

Best Local Similarity 50.0%; Pred. No. 5.2;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 IPVLDENGLFAP 12
      :|||:|
Db 7 VPALNENGFRYP 18
      :|||:|

RESULT 3
T35196
probable DNA gyrase chain B - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T35196
R:Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1998
A:Reference number: Z21571
A:Accession: T35196
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-707 <SES>
A:Cross-references: UNIPROT:O69998; EMBL:AL022374; PIDN:CAA18520.1; GSPDB:GN00070; SCOD
A:Experimental source: strain A3(2)
C:Genetics:
C:Supertfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 65.1%; Score 41; DB 2; Length 707;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDENGLFAP 12
      :|||:|
Db 290 VPVLDEDCQWTP 301
      :|||:|

RESULT 4
F69185
yhcv homolog MTH644 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: F69185
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: F69185
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-157 <MTH>
A:Cross-references: UNIPROT:O26740; GB:AE000844; GB:AE000666; NID:g2621707; PIDN:AA88514
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH644
C:Superfamily: conserved hypothetical protein yhcV; CBS homology
F:108-156/Domain: CBS homology <CBS>

Query Match 63.5%; Score 40; DB 1; Length 157;
Best Local Similarity 63.6%; Pred. No. 6.8;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLFA 11
      :|||:|
Db 133 LPVIDENGRLA 143
      :|||:|

RESULT 5
AF2724
hypothetical protein Atul201 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AF2724
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
```

```
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF2724
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <KUR>
A:Cross-references: UNIPROT:Q8UG41; GB:AE008688; PIDN:AA142212.1; PID:gl7739605; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atul201
A:Map position: circular chromosome

Query Match 63.5%; Score 40; DB 2; Length 337;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VLDENGLFAP 12
      :|||:|
Db 272 MLDETGLYAP 281
      :|||:|

RESULT 6
B97506
peptide ABC transporter, ATP-binding protein [imported] - Agrobacterium tumefaciens (str
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97506
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B97506
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <KUR>
A:Cross-references: UNIPROT:Q8UG41; GB:AE007869; PIDN:AAK87003.1; PID:gl5156245; GSPDB:G
C:Genetics:
A:Gene: AGR C. 2214
A:Map position: circular chromosome

Query Match 63.5%; Score 40; DB 2; Length 337;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VLDENGLFAP 12
      :|||:|
Db 272 MLDETGLYAP 281
      :|||:|

RESULT 7
H97350
two CBS domain containing protein [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: H97350
R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: H97350
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140 <KUR>
A:Cross-references: UNIPROT:Q97D09; GB:AE001437; PIDN:AAK81595.1; PID:gl5026776; GSPDB:G
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
```

A;Gene: CAC3674

Query Match 61.9%; Score 39; DB 2; Length 140;  
Best Local Similarity 60.0%; Pred. No. 9.1;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 10  
|||:|:|:|:  
Db 37 IPILDDNGKY 46

RESULT 8  
AB3238  
nitrotriacetate monooxygenase, component A Atu6084 [imported] - Agrobacterium tumefaciens  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: AB3238  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AB3238  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-450 <KUR>  
A;Cross-references: UNIPROT:Q8U651; GB:AE008690; PIDN:AAL46320.1; PID:g17744106; GSPDB:C  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atu6084  
A;Genome: plasmid  
C;Superfamily: nitrotriacetate monooxygenase

Query Match 61.9%; Score 39; DB 2; Length 450;  
Best Local Similarity 70.0%; Pred. No. 34;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 10  
|||:|:|:|:  
Db 406 VPVLQERGLF 415

RESULT 9  
G75318  
DNA ligase - Deinococcus radiodurans (strain R1)  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 12-Jul-2004  
C;Accession: G75318  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A;Reference number: A75250; MUID:20036896; PMID:10567266  
A;Accession: G75318  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-700 <WHI>  
A;Cross-references: UNIPROT:Q9RSQ5; GB:AE002043; GB:AE000513; NID:g6459859; PIDN:AAF1161  
A;Experimental source: strain R1  
C;Genetics:  
A;Gene: DR2069  
A;Map position: 1  
C;Superfamily: NAD+-dependent DNA ligase, LigA type

Query Match 61.9%; Score 39; DB 2; Length 700;  
Best Local Similarity 72.7%; Pred. No. 57;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLF 11

Db 661 VPVLDEAGLAA 671  
:|||||:|:|

RESULT 10  
S64951  
hypothetical protein YLR114c - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein L2941  
C;Species: Saccharomyces cerevisiae  
C;Date: 01-Aug-1995 #sequence\_revision 24-May-1996 #text\_change 09-Jul-2004  
C;Accession: S64951; S69401  
R;Verhasselt, P.; Voet, M.; Volckaert, G.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64943  
A;Accession: S64951  
A;Molecule type: DNA  
A;Residues: 1-764 <VER>  
A;Cross-references: UNIPROT:Q12500; EMBL:Z73286; NID:g1360509; PIDN:CAA97681.1; PID:e245  
A;Experimental source: strain S288C  
R;Verhasselt, P.; Volckaert, G.  
submitted to the EMBL Data Library, September 1995  
A;Reference number: S69393  
A;Accession: S69401  
A;Molecule type: DNA  
A;Residues: 1-764 <VEW>  
A;Cross-references: EMBL:X89514; NID:g1297019; PIDN:CAA61692.1; PID:e198747; PID:g129702  
C;Genetics:  
A;Gene: SGD:EPR4  
A;Cross-references: SGD:S0004104  
A;Map position: 12R

Query Match 61.9%; Score 39; DB 2; Length 764;  
Best Local Similarity 41.7%; Pred. No. 62;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFAP 12  
:|:|:|:|:|:|

Db 277 LPFKEKGLFSP 288

RESULT 11  
A82255  
hypothetical protein VC0998 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: A82255  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: A82255  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1621 <HEI>  
A;Cross-references: UNIPROT:Q9KTA5; GB:AE004181; GB:AE003852; NID:g9655454; PIDN:AAF9415  
A;Experimental source: serogroup O1; strain N16961; biotype E1 for  
C;Genetics:  
A;Gene: VC0998  
A;Map position: 1

Query Match 61.9%; Score 39; DB 2; Length 1621;  
Best Local Similarity 80.0%; Pred. No. 1.5e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PVLDEGLFA 11  
|||||

Db 743 PVLDENEAFA 752

RESULT 12  
A64556

```

toxin-like outer membrane protein HP0289 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: A64556
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen-
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 380, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520, MUID:97394467, PMID:9252185
A:Accession: A64556
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2893 <TOM>
A:Cross-references: UNIPROT:O25063; GB:AE000547; GB:AE000511; NID:g2313377; PIDN:AAD0735

Query Match 61.9%; Score 39; DB 2; Length 2893;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IPVLDENGLFAP 12
||| : |||||
Db 2077 IPNLGKKGIFAP 2088

RESULT 13
T08505
trbH protein - Enterobacter aerogenes plasmid R751
C:Species: Enterobacter aerogenes
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08505
R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
Plasmid 36, 95-111, 1996
A>Title: Conservation of the genetic switch between replication and transfer genes of In
A:Reference number: Z16434; MUID:97118926; PMID:8954881
A:Accession: T08505
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-162 <THO>
A:Cross-references: UNIPROT:P71183; EMBL:U67194; NID:g1572520; PIDN:AAC64449.1; PID:g157
C:Genetics:
A:Gene: trbH
A:Genome: plasmid R751

Query Match 60.3%; Score 38; DB 2; Length 162;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLDENGLFAP 12
||| : |||||
Db 144 PYLAQNGTFAP 154

RESULT 14
AC0929
hypochemical protein STY3692 [imported] - Salmonella enterica subsp. enterica serovar Ty
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0929
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Comerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0929
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD09453.1; PID:g16504570; GSPDB:GN00176

```

```

C:Genetics:
A:Gene: STY3692
C:Superfamily: phage T4 tail fiber assembly protein gp38

Query Match 60.3%; Score 38; DB 2; Length 197;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PVLDENGL 9
||| : |||||
Db 3 PVFDENGL 10

RESULT 15
A90664
hypochemical protein EC80281 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A90664
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A90664
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <HAY>
A:Cross-references: UNIPROT:Q8X7M0; GB:BA000007; PIDN:BA833704.1; PID:g13359738; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC80281
C:Superfamily: phage T4 tail fiber assembly protein gp38

Query Match 60.3%; Score 38; DB 2; Length 197;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PVLDENGL 9
||| : |||||
Db 3 PVFDENGL 10

RESULT 16
F85514
unknown protein from prophage CP-933H [imported] - Escherichia coli (strain O157:H7, sub:
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F85514
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85514
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <STO>
A:Cross-references: UNIPROT:Q8X7M0; GB:AE005174; NID:g12513030; PIDN:AAG54578.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL593
C:Genetics:
A:Gene: Z0316
C:Superfamily: phage T4 tail fiber assembly protein gp38

Query Match 60.3%; Score 38; DB 2; Length 197;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PVLDENGL 9
||| : |||||
Db 3 PVFDENGL 10

```

RESULT 17  
G69623  
flagellar hook-basal body protein flhP - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: G69623; I40533  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koeter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, K.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: G69623  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-268 <UN>  
A;Cross-references: UNIPROT:P39753; GB:299122; GB:AL009126; NID:G2636029; PIDN:CAB15656.  
A;Experimental source: strain 168  
R;Abhayawardhane, Y.; Stewart, G.C.  
J. Bacteriol. 177, 765-773, 1995  
A;Title: Bacillus subtilis possesses a second determinant with extensive sequence simila  
A;Reference number: I40531; MUID:95138040; PMID:7836311  
A;Accession: I40533  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-67, 'GLT', '71', 'G', '72-134 <RES>  
A;Cross-references: EMBL:U12962; NID:G531459; PIDN:AAA67880.1; PID:G531462  
C;Genetics:  
A;Gene: flhP  
C;Superfamily: rod protein flgP  
Query Match 60.3%; Score 38; DB 2; Length 268;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 PVLDENG 8  
Db 142 PVLDENG 148  
RESULT 18  
E97298  
uncharacterized Fe-S protein, pflX (pyruvate formate lyase activating protein) homolog  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: E97298  
R;Nolling, J.; Bretton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clc  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: E97298  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-298 <KUR>  
A;Cross-references: UNIPROT:Q97E74; GB:AB001437; PIDN:AAK61176.1; PID:G15026315; GSPDB:C  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC3242  
C;Superfamily: conserved hypothetical protein MJ0674  
Query Match 60.3%; Score 38; DB 2; Length 298;  
Best Local Similarity 87.5%; Pred. No. 33;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 PVLDENG 9  
Db 192 PVLDENG 199  
RESULT 19  
RGBYC3  
regulatory protein SNF4 - yeast (Saccharomyces cerevisiae)  
N;Alternate names: CAT3 protein; protein G2945; protein YGL115W  
C;Species: Saccharomyces cerevisiae  
C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
C;Accession: A38906; J70316; S48508; S64125  
R;Ceizenza, J.L.; Eng, F.J.; Carlson, M.  
Mol. Cell. Biol. 9, 5045-5054, 1989  
A;Title: Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae: evidence for p  
A;Reference number: A33480; MUID:90097921; PMID:2481228  
A;Accession: A38906  
A;Molecule type: DNA  
A;Residues: 1-322 <CEL>  
A;Cross-references: UNIPROT:P12904; GB:M30470; NID:G172635; PIDN:AAA35061.1; PID:G172635  
R;Schueller, H.J.; Entian, K.D.  
Gene 67, 247-257, 1988  
A;Title: Molecular characterization of yeast regulatory gene CAT3 necessary for glucose  
A;Reference number: J70316; MUID:89006284; PMID:3049255  
A;Accession: J70316  
A;Molecule type: DNA  
A;Residues: 1-322 <SCH>  
A;Cross-references: GB:M21760; NID:G171164; PIDN:AAA34472.1; PID:G171165  
R;Doi, A.; Doi, K.  
submitted to the EMBL Data Library, June 1993  
A;Description: Correct end of the ORF for the CDC20 gene of Saccharomyces cerevisiae.  
A;Reference number: S48507  
A;Accession: S48508  
A;Molecule type: DNA  
A;Residues: 1-21 <DOI>  
A;Cross-references: EMBL:D16506; NID:G391938; PIDN:BAA03958.1; PID:G2160324  
R;Lauquin, G.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64122  
A;Accession: S64125  
A;Molecule type: DNA  
A;Residues: 1-322 <LAU>  
A;Cross-references: EMBL:Z72637; NID:G1322666; PIDN:CAA96923.1; PID:G1322667; GSPDB:GN00  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: SGD:SNF4; CAT34; MIPS:YGL115W  
A;Cross-references: SGD:S0003083; MIPS:YGL115W  
C;Map position: 7L  
C;Function:  
A;Description: involved in derepression of glucose-repressed genes  
C;Superfamily: CAT3 protein  
C;Keywords: nucleus; transcription regulation  
Query Match 60.3%; Score 38; DB 1; Length 322;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 I PVLDENG 8  
Db 223 VPIIDENG 230  
RESULT 20  
AF0864  
membrane-bound lytic murein transglycosylase A precursor STY3128 [imported] - Salmonella  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AF0864  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AF0864  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-365 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD02814.1; PID:g16504068; GSPDB:GN00176  
C;Genetics:  
A;Gene: STY3128

Query Match 60.3%; Score 38; DB 2; Length 365;  
Best Local Similarity 54.5%; Pred. No. 41;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLPA 11  
:|||||:  
Db 295 VPVLDENGKFS 305  
:|||||:

RESULT 21  
D84618  
probable amino acid acetyltransferase [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Aug-2004  
C;Accession: D84618  
R;Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eise, J.A.; Salzb, S.D.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzb, S.D.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: D84618  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-620 <STO>  
A;Cross-references: GB:AE002093; NID:g3445208; PIDN:AAC32438.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g22910  
A;Map position: 2  
C;Superfamily: Amino-acid N-acetyltransferase with amino acid kinase domain

Query Match 60.3%; Score 38; DB 2; Length 620;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVLDENG 8  
:|||||:  
Db 300 PVLDENG 306  
:|||||:

RESULT 22  
E82221  
DNA gyrase, chain A VC1258 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: E82221  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: E82221  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-894 <HEI>  
A;Cross-references: UNIPROT:Q9KJ8; GB:AE004205; GB:AE003852; NID:g9655740; PIDN:AAF9441  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC1258

A;Map position: 1  
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (atrac)  
Query Match 60.3%; Score 38; DB 2; Length 894;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVLDENG 8  
:|||||:  
Db 720 PVLDENG 726  
:|||||:

RESULT 23  
G95057  
endo-beta-N-acetylglucosaminidase, probable [imported] - Streptococcus pneumoniae (strain  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: G95057  
R;Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eise, J.A.; Salzb, S.D.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzb, S.D.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: D84618  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-620 <STO>  
A;Cross-references: GB:AE002093; NID:g3445208; PIDN:AAC32438.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g22910  
A;Map position: 2  
C;Superfamily: Amino-acid N-acetyltransferase with amino acid kinase domain

Query Match 59.5%; Score 37.5; DB 2; Length 1659;  
Best Local Similarity 52.9%; Pred. No. 2.8e+02;  
Matches 9; Conservative 1; Mismatches 0; Indels 7; Gaps 1;

QY 3 VLDEN-----GLFAP 12  
:|||||:  
Db 457 ILDENGKRLSLGLFAP 473  
:|||||:

RESULT 24  
H97926  
hypothetical protein spr0440 [imported] - Streptococcus pneumoniae (strain R6)  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: H97926  
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E.  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
J. P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A;Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: H97926  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1659 <KUR>  
A;Cross-references: UNIPROT:Q8CZ52; GB:AE007317; PIDN:AAK99244.1; PID:g15458007; GSPDB:G  
C;Genetics:  
A;Gene: spr0440

Query Match 59.5%; Score 37.5; DB 2; Length 1659;  
Best Local Similarity 52.9%; Pred. No. 2.8e+02;  
Matches 9; Conservative 1; Mismatches 0; Indels 7; Gaps 1;

QY 3 VLDEN-----GLFAP 12  
:|||||:  
Db 457 ILDENGKRLSLGLFAP 473  
:|||||:

```
RESULT 25
F83988
inosine-5'-monophosphate dehydrogenase BH2710 [imported] - Bacillus halodurans (strain C
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: F83988
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F83988
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-144 <STO>
A;Cross-references: UNIPROT:Q9X9D8; GB:AP001516; GB:BA000004; NID:gl0175192; PIDN:BA064
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2710

Query Match 58.7%; Score 37; DB 2; Length 144;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDEGLF 10
Db 106 VPVDDGIF 115
||:|:|:|:|

RESULT 26
AB0982
Probable 2-hydroxyacid dehydrogenase STY4156 [imported] - Salmonella enterica subsp. ent
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AB0982
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AB0982
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD07982.1; PID:g16504968; GSPDB:GN00176
C;Genetics:
A;Gene: STY4156
C;Superfamily: phosphoglycerate dehydrogenase

Query Match 58.7%; Score 37; DB 2; Length 324;
Best Local Similarity 70.0%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLDEGLFPA 11
Db 239 PVDENALIA 248
||:|:|:|:|

RESULT 27
C65154
Probable 2-hydroxyacid dehydrogenase in bisc-cepA intergenic region - Escherichia coli (
N;Alternate names: hypothetical protein o365
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: C65154; S47774
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C65154
```

```
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-328 <BLAT>
A;Cross-references: GB:AE000432; GB:U00096; NID:G2367241; PIDN:AAC76577.1; PID:G2367243;
A;Experimental source: strain K-12, substrain MG1655
R;Plunkett, G.
submitted to the EMBL Data Library, March 1994
A;Reference number: S47666
A;Accession: S47774
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327,'TKPRLRRVKPNRSGFYSLLGCNSCCSPGLIKRLMLCR' <PLU>
A;Cross-references: EMBL:U00039; NID:g466582; PIDN:AAB18530.1; PID:g466691
A;Note: this sequence has been corrected
C;Genetics:
A;Gene: YiaB
C;Superfamily: phosphoglycerate dehydrogenase

Query Match 58.7%; Score 37; DB 2; Length 328;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLDEGLFPA 11
Db 243 PVDENALIA 252
||:|:|:|:|

RESULT 28
F91183
Probable dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 050
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F91183
R;Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasaawata, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91183
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-328 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA037861.1; PID:gl33363912; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC84438
C;Superfamily: phosphoglycerate dehydrogenase

Query Match 58.7%; Score 37; DB 2; Length 328;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLDEGLFPA 11
Db 243 PVDENALIA 252
||:|:|:|:|

RESULT 29
B86030
Probable dehydrogenase viaE [imported] - Escherichia coli (strain O157:H7, substrain EDL
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: B86030
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B86030
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-328 <STO>
A;Cross-references: GB:AE005174; NID:gl2518288; PIDN:AAG58702.1; GSPDB:GN00145; UWGP:Z49
```

A;Experimental source: strain O157:H7, substrain EDJ933  
C;Genetics:  
A;Gene: ylaE  
C;Superfamily: phosphoglycerate dehydrogenase

Query Match 58.7%; Score 37; DB 2; Length 328;  
Best Local Similarity 70.0%; Pred. No. 56;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLDENGLFA 11  
||:|||||  
Db 243 PVDENALIA 252

RESULT 30  
AG2251  
translation initiation factor IP-2B alpha chain [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AG2251  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A;Reference number: ABL807; MUID:21595285; PMID:11759840  
A;Accession: AG2251  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-347 <KUR>  
A;Cross-references: UNIPROT:O8YR82; GB:BA000019; PIDN:BAB75265.1; PID:gl7132699; GSPDB:G000000000  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all3566  
C;Superfamily: translation initiation factor eIF-2B

Query Match 58.7%; Score 37; DB 2; Length 347;  
Best Local Similarity 60.0%; Pred. No. 60;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VLDENGLPAP 12  
:::|||||  
Db 326 IITENGAFAP 335

RESULT 31  
I55210  
tricarboxylate carrier - rat (fragment)  
C;Species: Rattus sp. (rat)  
C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 26-May-2000  
C;Accession: I55210  
R;Azzi, A.; Glerum, M.; Koller, R.; Mertens, W.; Spycher, S. J. Bioenerg. Biomembr. 25, 515-524, 1993  
A;Title: The mitochondrial tricarboxylate carrier.  
A;Reference number: I55210; MUID:94179133; PMID:81132491  
A;Accession: I55210  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-357 <RES>  
A;Cross-references: GB:S70011; NID:gs45997; PIDN:AA30258.1; PID:gs45998  
C;Superfamily: Saccharomyces probable membrane protein YOR271c

Query Match 58.7%; Score 37; DB 2; Length 357;  
Best Local Similarity 87.5%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDENG 8  
|||:|||||  
Db 240 IPVTDENG 247

RESULT 32  
C85933

membrane-bound lytic murein transglycosylase A [imported] - Escherichia coli (strain O157:H7)  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: C85933  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: C85933  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-365 <STO>  
A;Cross-references: UNIPROT:P46885; GB:AE005174; NID:gl2517295; PIDN:AAG57927.1; GSPDB:G000000000  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: mlcA

Query Match 58.7%; Score 37; DB 2; Length 365;  
Best Local Similarity 60.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLF 10  
:::|||||  
Db 295 VPLLDNNGKF 304

RESULT 33  
A98088

membrane-bound lytic murein transglycosylase A EC3673 [imported] - Escherichia coli (strain O157:H7)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: A98088  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome analysis.  
A;Reference number: A98088; MUID:21156231; PMID:11258796  
A;Accession: A98088  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-365 <HAY>  
A;Cross-references: UNIPROT:P46885; GB:BA000007; PIDN:BA337096.1; PID:gl3363145; GSPDB:G000000000  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: EC3673

Query Match 58.7%; Score 37; DB 2; Length 365;  
Best Local Similarity 60.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDENGLF 10  
:::|||||  
Db 295 VPLLDNNGKF 304

RESULT 34  
A65064

membrane-bound lytic murein transglycosylase A (EC 3.2.1.-) precursor - Escherichia coli (strain O157:H7)  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C;Accession: A65064  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co. A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A65064; MUID:97426617; PMID:9278503  
A;Accession: A65064  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-365 <BLAT>  
A;Cross-references: UNIPROT:P46885; GB:AE000365; GB:U00096; NID:g2367163; PIDN:AAC75855.1  
A;Experimental source: strain K-12, substrain MG1655

```
C;Genetics:
A;Gene: mlcA
C;Keywords: glycosidase; hydrolase

Query Match      58.7%; Score 37; DB 2; Length 365;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IPVLDEGLF 10
   :||| |||
Db 295 VPLDNNNGK 304

RESULT 35
T44324
hypothetical protein wbf1 [imported] - Vibrio cholerae
C;Species: Vibrio cholerae
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T44324
R;Yamaoka, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda, Y.
Gene 237, 321-332, 1999
A;Title: The genes responsible for O-antigen synthesis of Vibrio cholerae O139 are close
A;Reference number: 222749; MUID:99453293; PMID:10521656
A;Accession: T44324
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-390 <YAM>
A;Cross-references: UNIPROT:O87153; EMBL:AB012957; NID:g4115688; PIDN:BA033628.1; PID:g3
A;Experimental source: strain O22
C;Genetics:
A;Note: wbf1
C;Superfamily: erythromycin resistance protein

Query Match      58.7%; Score 37; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LDENGLF 10
   |||||
Db 364 LDENGLF 370

RESULT 36
AC2638
acyl-CoA thiolase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AC2638
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erge, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <KUR>
A;Cross-references: UNIPROT:Q8UH28; GB:AE008688; PIDN:AAL41521.1; PID:gl7738851; GSPDB:G
C;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu0502
A;Map position: circular chromosome
C;Superfamily: acetyl-CoA acetyltransferase

Query Match      58.7%; Score 37; DB 2; Length 402;
Best Local Similarity 77.8%; Pred. No. 71;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDEGL 9
   ||| |||
Db 191 IPVKDQNG 199

C;Genetics:
A;Gene: mlcA
C;Keywords: glycosidase; hydrolase

Query Match      58.7%; Score 37; DB 2; Length 402;
Best Local Similarity 77.8%; Pred. No. 71;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDEGL 9
   ||| |||
Db 191 IPVKDQNG 199

RESULT 37
B97420
probable acyl-CoA thiolase (PA1736) [imported] - Agrobacterium tumefaciens (strain C58,
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: B97420
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97420
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <KUR>
A;Cross-references: UNIPROT:Q8UH28; GB:AE007869; PIDN:AAK86315.1; PID:gl5155433; GSPDB:G
C;Genetics:
A;Gene: AGR_C_886
A;Map position: circular chromosome
C;Superfamily: acetyl-CoA acetyltransferase

Query Match      58.7%; Score 37; DB 2; Length 402;
Best Local Similarity 77.8%; Pred. No. 71;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDEGL 9
   ||| |||
Db 191 IPVKDQNG 199

RESULT 38
C85140
serine/threonine-specific protein kinase MHK [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Aug-2004
C;Accession: C85140
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: C85140
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-443 <STO>
A;Cross-references: UNIPROT:P43294; GB:NC_001268; NID:g7268004; PIDN:CAB78344.1; GSPDB:G
C;Genetics:
A;Gene: Atg13020
A;Map position: 4
C;Superfamily: protein kinase homology

Query Match      58.7%; Score 37; DB 2; Length 443;
Best Local Similarity 87.5%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 DENGFLAP 12
   ||||| |||
Db 382 DENGFLAP 389

RESULT 39
S38327
serine/threonine-specific protein kinase MHK (EC 2.7.1.-) - Arabidopsis thaliana
N;Alternate names: mak homologous kinase; protein P25G13.110
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Aug-2004
C;Accession: S38327; T10204
R;Moran, T.V.; Walker, J.C.
Biochim. Biophys. Acta 1216, 9-14, 1993
A;Title: Molecular cloning of two novel protein kinase genes from Arabidopsis thaliana.
A;Reference number: S38326; MUID:94032493; PMID:8218420
```

A;Accession: S38327  
A;Molecule type: mRNA  
A;Residues: 1-443 <MOR>  
A;Cross-references: UNIPROT:P43294; EMBL:L07249; NID:g166810; PIDN:AAA18854.1; PID:g166810  
R;Bevan, M.; Pohl, T.; Weizensegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16991  
A;Accession: T10204

A;Molecule type: DNA  
A;Residues: 1-65, 'N', 67-443 <BEV>  
A;Cross-references: EMBL:AL079349; GSPDB:GN00062; ATSP:F25G13.110  
A;Experimental source: cultivar Columbia; BAC clone F25G13  
C;Genetics:  
A;Gene: MHK; ATSP:F25G13.110  
A;Map position: 4  
A;Introns: 37/3; 60/3; 137/1; 171/2; 189/1; 211/2; 247/3; 269/3; 292/3; 312/1; 347/1  
C;Superfamily: protein kinase homology  
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F10-267/Domain: protein kinase homology <KIN>  
F18-26/Region: protein kinase ATP-binding motif

Query Match 58.7%; Score 37; DB 2; Length 443;  
Best Local Similarity 87.5%; Pred. No. 79;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 DENGFLAP 12  
||| ||| ||  
Db 382 DENGFLAP 389

RESULT 40  
T19628  
hypothetical protein C32A3.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19628  
R;Thomas, K.  
submitted to the EMBL Data Library, February 1995  
A;Reference number: Z19154  
A;Accession: T19628  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-468 <WIL>  
A;Cross-references: UNIPROT:Q09260; EMBL:Z48241; PIDN:CAA88284.1; GSPDB:GN000021; CESP:C32A3  
A;Experimental source: Clone C32A3  
C;Genetics:  
A;Gene: CESP:C32A3.1  
A;Map position: 3  
A;Introns: 33/2; 155/3; 261/3; 409/2

Query Match 58.7%; Score 37; DB 2; Length 468;  
Best Local Similarity 54.5%; Pred. No. 84;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLDENGFLAP 12  
|:|:|:|:|  
Db 238 PVDENNLAVP 248

Search completed: January 7, 2005, 10:06:00  
Job time : 27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2005, 10:01:57 ; Search time 108 Seconds  
(without alignments)  
63.931 Million cell updates/sec

Title: US-09-699-224A-1  
Perfect score: 63  
Sequence: 1 IPLVDENGLFAP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	73.0	332	2	Q96ZS5 sulfolobus
2	46	73.0	366	1	MLTA_BUCAP
3	45	71.4	359	1	P57531 buchnera ap
4	43	68.3	136	2	Q8XIG7
5	43	68.3	955	2	Q89VU0
6	42	66.7	942	2	Q6S734
7	42	66.7	942	2	RA555735
8	42	66.7	1058	2	Q6XPT1
9	42	66.7	1058	2	CAE75591
10	42	66.7	1058	2	AAP46169
11	42	66.7	1576	2	Q8CM69
12	41	65.1	180	2	Q7CS40
13	41	65.1	209	2	Q6XIF9
14	41	65.1	209	2	AAR09894
15	41	65.1	245	2	O31394
16	41	65.1	320	2	Q6FMK9
17	41	65.1	521	2	Q728L3
18	41	65.1	521	2	RA597062
19	41	65.1	696	2	Q7VRX7
20	41	65.1	696	2	Q7W0T4
21	41	65.1	696	2	Q7WCJ7
22	41	65.1	702	2	Q83VM6
23	41	65.1	702	2	Q83WB8
24	41	65.1	707	2	Q69998
25	41	65.1	778	2	Q6CPB9
26	41	65.1	2255	2	Q71TF8
27	41	65.1	2255	2	AAQ13985
28	41	65.1	2255	2	AAQ14093
29	41	65.1	2462	2	Q8RGZ3
30	41	65.1	2806	2	Q8RI19
31	41	65.1	3119	2	Q7P6S5

32 40 63.5 157 2 Q26740 methanobact  
33 40 63.5 224 2 Q9V8G6 drosophila  
34 40 63.5 233 2 Q7PSW6 anopheles g  
35 40 63.5 233 2 Q8MUQ6 anopheles g  
36 40 63.5 236 2 Q6YUW2 oryza sativ  
37 40 63.5 236 2 BAD16458  
38 40 63.5 333 2 Q841U3 vibrio vuln  
39 40 63.5 333 2 Q7MI83 vibrio vuln  
40 40 63.5 333 2 Q8DBN6 vibrio vuln  
41 40 63.5 337 2 Q7CZS0 agrobacteri  
42 40 63.5 337 2 Q8UG41 agrobacteri  
43 40 63.5 365 2 Q83A93 coxiella bu  
44 40 63.5 376 2 Q8A361 bacteroides  
45 40 63.5 397 2 Q7N2P9 photorhabdu

## ALIGNMENTS

## RESULT 1

Q96ZS5 PRELIMINARY; PRT; 332 AA.  
AC Q96ZS5;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein ST1760.  
GN OrderedLocuNames=ST1760;  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI\_TaxID=111955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JCM 10545 / 7;  
RX MEDLINE=21456156; PubMed=11572479;  
RA Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermoacidophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain 7.";  
RL DNA Res. 8:123-140(2001).  
DR EMBL; AP000987; BAB66848.1; -.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 332 AA; 38801 MW; ACAD1F64C3AFEEA9 CRC64;

Query Match 73.0%; Score 46; DB 2; Length 332;  
Best Local Similarity 66.7%; Pred. No. 6.4;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IPLVDENGLFAP 12  
Db 161 IPLVDENGLIMP 172  
|:::|||||  
|

## RESULT 2

MLTA\_BUCAP STANDARD; PRT; 366 AA.  
ID MLTA\_BUCAP  
AC Q8K9A7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Membrane-bound lytic murein transglycosylase A homolog (EC 3.2.1.-)  
DE (Murein hydrolase A).  
GN Name=mltA; OrderedLocuNames=BUG9442;  
OS Buchnera aphidicola (subsp. Schizaphis graminum).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Buchnera.  
OX NCBI\_TaxID=98794;  
RN [1]



```

Query Match          66.7%; Score 42; DB 2; Length 942;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1  IPVLDENGLFPA 11
      :||:| |||:
DB      910  LPVIDSNGIFS 920

RESULT 7
AAS55735
ID  AAS55735      PRELIMINARY;      PRT;      942 AA.
AC  AAS55735:
DT  01-JUN-2004 (TEMBLrel. 27, Created)
DT  01-JUN-2004 (TEMBLrel. 27, Last sequence update)
DT  01-JUN-2004 (TEMBLrel. 27, Last annotation update)
DE  Aryl hydrocarbon receptor 2b (Fragment).
DE  AHR2.
GN  AHR2.
OS  Salmo salar (Atlantic salmon).
OS  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
NCBI_TaxID=8030;
RN  [1]
RP  SEQUENCE FROM N.A.
RT  Hansson M., Wittzell H., Persson K., Von Schantz T.;
RT  "Unprecedented genomic diversity of AhR1 and AHR2 genes in Atlantic
RT  salmon (Salmo salar L.).";
RT  Aquat. Toxicol. 0:0-0(2004).
DR  EMBL; AY463929; AAS55735.1; .
DR  EMBL; AY463927; AAS55735.1; JOINED.
DR  EMBL; AY463928; AAS55735.1; JOINED.
KW  Receptor.
FT  NON TER
FT  NON TER
SQ  SEQUENCE      942 AA; 102536 MW; 4755C0FD7D8AD7EE CRC64;

Query Match          66.7%; Score 42; DB 2; Length 942;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1  IPVLDENGLFPA 11
      :||:| |||:
DB      910  LPVIDSNGIFS 920

RESULT 8
Q6XPT1
ID  Q6XPT1      PRELIMINARY;      PRT;      1058 AA.
AC  Q6XPT1;
DT  05-JUL-2004 (TEMBLrel. 27, Created)
DT  05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE  Aryl hydrocarbon receptor 2 beta.
DE  Name=ahr2b;
GN  Ahr2b.
OS  Salmo salar (Atlantic salmon).
OS  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
NCBI_TaxID=8030;
RN  [1]
RP  SEQUENCE FROM N.A.
RT  Hansson M., Wittzell H., Persson K., Von Schantz T.;
RT  "Two additional Ahr2 genes in Atlantic salmon.";
RT  Aquat. Toxicol. 0:0-0(2004).
RN  [2]
RP  SEQUENCE FROM N.A.
RT  Hansson M.C., Persson K.;
RT  Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RT  TISSUE=Testis;
RT  Bermanian V., Ladstein S., Goksoyr A., Male R.;

```

```

RP SEQUENCE FROM N.A.
RA Hanson M., Wittzell H., Persson K., Von Schantz T.;
RT "Two additional AHR2 genes in Atlantic salmon.";
RL Aquat. Toxicol. 0:0-0(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Hanson M.C., Persson K.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RW EMBL; AY219865; AAP46169.1; -.
KW Receptor.
SQ SEQUENCE 1058 AA; 115396 MW; AFD7FE2D1BA91061 CRC64;

Query Match 66.7%; Score 42; DB 2; Length 1058;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Indels

QY 1 IPVLDENGLEA 11
      |||:|||:|:|
DB 1026 LPVIDSNGIFS 1036

RESULT 11
Q8CM69 PRELIMINARY; PRT; 1576 AA.
AC Q8CM69;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein gbs0386 (Hypothetical protein gbs0716)
DE DE Hypothetical protein gbs0993.
DE OrderedLocusNames=gbs0386 gbs0716, gbs0993;
OS Streptococcus agalactiae NEM316.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus; Streptococcus agalactiae serogroup III.
ON NCBI_TaxID=211110;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangeul
RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen cau
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766845; CAD46030.1; -.
DR EMBL; AL766846; CAD46360.1; -.
DR EMBL; AL766848; CAD46652.1; -.
DR SagalList; gbs0386; -.
DR SagalList; gbs0716; -.
DR SagalList; gbs0993; -.
DR InterPro; IPR010359; DUF955.
DR Pfam; PF06114; DUF955; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 1576 AA; 179786 MW; 8F1D40311C7BB3E8 CRC64;

Query Match 66.7%; Score 42; DB 2; Length 1576;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 0; Indels

QY 1 IPVLDENG 8
      |||:|||:|
DB 1315 IPVLDENG 1322

RESULT 12
Q7CS40 PRELIMINARY; PRT; 180 AA.
AC Q7CS40;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR L 3039D.

```

```

GN OrderedLocusNames=AGR_L_3039;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA William C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,
RA Cleo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR ENBL; AE008353; AAK90091.1; -.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 1.
SQ SEQUENCE 180 AA; 19735 MW; 71D0FF440AE034F0 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 180;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDENGLPAP 12
Db 7 VPALNENGFYRP 18

RESULT 13
Q6XIF9 PRELIMINARY; PRT; 209 AA.
ID Q6XIF9
AC Q6XIF9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to Drosophila melanogaster CG17524 (Fragment).
OS Drosophila yakuba (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7245;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22887302; PubMed=14525923;
RA Domazet-Lozo T., Tautz D.;
RT "An evolutionary analysis of orphan genes in Drosophila.";
RL Genome Res. 13:2213-2219(2003).
DR ENBL; AY231871; AAR09894.1; -.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C-like.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
FT NON_TER 1
FT NON_TER 209
SQ SEQUENCE 209 AA; 23559 MW; 5D6A8D4F6F13DC5C CRC64;

Query Match 65.1%; Score 41; DB 2; Length 209;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDENGLP 10
Db 55 VPVLDDNGFY 64

RESULT 14
AAR09894 PRELIMINARY; PRT; 209 AA.
ID AAR09894
AC AAR09894
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to Drosophila melanogaster CG17524 (Fragment).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7245;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22887302; PubMed=14525923;
RA Domazet-Lozo T., Tautz D.;
RT "An evolutionary analysis of orphan genes in Drosophila.";
RL Genome Res. 13:2213-2219(2003).
DR ENBL; AY231871; AAR09894.1; -.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C-like.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
FT NON_TER 1
FT NON_TER 209
SQ SEQUENCE 209 AA; 23559 MW; 5D6A8D4F6F13DC5C CRC64;

Query Match 65.1%; Score 41; DB 2; Length 209;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDENGLP 10
Db 55 VPVLDDNGFY 64

RESULT 15
Q31394 PRELIMINARY; PRT; 245 AA.
ID Q31394
AC Q31394
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sensory transduction histidine kinase.
GN Namesorf2;
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RA Vazquez G.J., Pettinari M.J., Mendez B.M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Mendez B.S.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR ENBL; Y14588; CAA74928.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR00989; PAS; 1.
DR SMART; SM00091; PAS; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS50112; PAS; 1.
KW Kinase.
SQ SEQUENCE 245 AA; 27391 MW; 6B6E2A39BB25E7C1 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 245;
Best Local Similarity 87.5%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPVLDENG 8
Db 199 IPILDENG 206

RESULT 16

```

```

Q6FMK9
ID Q6FMK9 PRELIMINARY; PRT; 320 AA.
AC Q6FMK9;
DR TIGR; DVU2590; -.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P12904|Saccharomyces cerevisiae YGL115w SNF4.
GN ORFNames=CAGL0K07161g;
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Tallia E.,
RA Goffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisarame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hanttraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in Yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380957; CAG61496.1; -.
DR InterPro; IPR00644; CBS.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
SQ SEQUENCE 320 AA; 36478 MW; EB736BAE06FF5737 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 320;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVLDENGCL 9
DB 223 VPILDENGV 231

RESULT 17
Q728L3 PRELIMINARY; PRT; 521 AA.
AC Q728L3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Sensory box protein.
GN OrderedLocusNames=DVU2590;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017318; AAS97062.1; -.
DR TIGR; DVU2590; -.
SQ SEQUENCE 521 AA; 57195 MW; 837255C37877532C CRC64;

Query Match 65.1%; Score 41; DB 2; Length 521;
Best Local Similarity 63.6%; Pred. No. 90;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PVLDENGLFAP 12
DB 269 PMLEENGIEAP 279

RESULT 19
Q7VRX7 PRELIMINARY; PRT; 696 AA.
AC Q7VRX7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA ligase [EC 6.5.1.2].
GN Name=liga; Synonyms=dnaL, lig, lop, pdeC; OrderedLocusNames=BP3560;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -|- FUNCTION: This protein catalyzes the formation of phosphodiester
CC linkages between 5'-phosphoryl and 3'-hydroxyl groups in double-
CC stranded DNA using NAD as a coenzyme and as the energy source for
CC the reaction. It is essential for DNA replication and repair of
CC damaged DNA (By similarity).
CC -|- CATALYTIC ACTIVITY: NAD(+) + {deoxyribonucleotide}(N) +
CC {deoxyribonucleotide}(M) = AMP + nicotinamide nucleotide +
CC {deoxyribonucleotide}(N+M).
CC -|- SIMILARITY: Belongs to the NAD-dependent DNA ligase family.
CC EMBL: BX640421; CA43819.1; -.
CC DR GO: GO:0005622; C:intracellular; IEA.
CC DR GO: GO:0003677; F:DNA binding; IEA.
CC DR GO: GO:0003911; F:DNA ligase (NAD+) activity; IEA.
CC DR GO: GO:0016874; F:ligase activity; IEA.
CC DR GO: GO:0006281; P:DNA repair; IEA.
CC DR GO: GO:0006260; P:DNA replication; IEA.
CC DR InterPro: IPR001357; BRCT.
CC DR InterPro: IPR004150; DNA_ligase_OB.
CC DR InterPro: IPR001679; DNALigase.
CC DR InterPro: IPR003583; HHH.
CC DR InterPro: IPR001357; BRCT.
CC DR InterPro: IPR004149; Znf_DNALigase_C4.
CC DR Pfam: PF00533; BRCT; 1.
CC DR SMART: SM00292; BRCT; 1.
CC DR SMART: SM00278; HHH1; 4.
CC DR Pfam: PF01120; DNA_ligase_OB; 1.
CC DR Pfam: PF03119; DNA_ligase_ZBD; 1.
CC DR Pfam: PF00633; HHH; 3.
CC DR ProDom: PD003944; DNALigase; 1.
CC DR SMART: SM00292; BRCT; 1.
CC DR SMART: SM00278; HHH1; 4.
CC DR TIGRFAMs: TIGR00575; dnlj; 1.
CC DR PROSITE: PS01053; LIGANC; 1.
CC DR PROSITE: PS01072; BRCT; 1.
CC DR PROSITE: PS01055; DNA_LIGASE_N1; 1.
CC DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
CC DR Complete proteome; DNA repair; DNA replication; Ligase; NAD.
SQ SEQUENCE 696 AA; 75585 MW; 2FB4177471837EA3 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 696;
Best Local Similarity 72.7%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFPA 11
Db 682 VPVLDEGLKA 692

RESULT 20
Q7W0T4 PRELIMINARY; PRT; 696 AA.
AC Q7W0T4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA ligase (EC 6.5.1.2).
GN Name=liga; Synonyms=dnlai, lig, lop, pdeC; OrderedLocusNames=BPP3353;
OS Bordetella parapertussis.

```

```

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=12822 / ATCC BAA-587;
RC MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -|- FUNCTION: This protein catalyzes the formation of phosphodiester
CC linkages between 5'-phosphoryl and 3'-hydroxyl groups in double-
CC stranded DNA using NAD as a coenzyme and as the energy source for
CC the reaction. It is essential for DNA replication and repair of
CC damaged DNA (By similarity).
CC -|- CATALYTIC ACTIVITY: NAD(+) + {deoxyribonucleotide}(N) +
CC {deoxyribonucleotide}(M) = AMP + nicotinamide nucleotide +
CC {deoxyribonucleotide}(N+M).
CC -|- SIMILARITY: Belongs to the NAD-dependent DNA ligase family.
CC EMBL: BX640433; CAE38638.1; -.
CC DR GO: GO:0005622; C:intracellular; IEA.
CC DR GO: GO:0003677; F:DNA binding; IEA.
CC DR GO: GO:0003911; F:DNA ligase (NAD+) activity; IEA.
CC DR GO: GO:0016874; F:ligase activity; IEA.
CC DR GO: GO:0006281; P:DNA repair; IEA.
CC DR GO: GO:0006260; P:DNA replication; IEA.
CC DR InterPro: IPR001357; BRCT.
CC DR InterPro: IPR001679; DNALigase.
CC DR InterPro: IPR004150; DNA_ligase_OB.
CC DR InterPro: IPR000445; Hhh.
CC DR InterPro: IPR003583; HHH.
CC DR InterPro: IPR004149; Znf_DNALigase_C4.
CC DR Pfam: PF00533; BRCT; 1.
CC DR Pfam: PF01653; DNA_ligase_aden; 1.
CC DR Pfam: PF03120; DNA_ligase_OB; 1.
CC DR Pfam: PF03119; DNA_ligase_ZBD; 1.
CC DR Pfam: PF00633; HHH; 3.
CC DR ProDom: PD003944; DNALigase; 1.
CC DR SMART: SM00292; BRCT; 1.
CC DR SMART: SM00278; HHH1; 4.
CC DR SMART: SM00532; LIGANC; 1.
CC DR TIGRFAMs: TIGR00575; dnlj; 1.
CC DR PROSITE: PS01072; BRCT; 1.
CC DR PROSITE: PS01055; DNA_LIGASE_N1; 1.
CC DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
CC DR Complete proteome; DNA repair; DNA replication; Ligase; NAD.
KW Complete proteome; DNA repair; DNA replication; Ligase; NAD.
SQ SEQUENCE 696 AA; 75502 MW; 9AB412B7D5406FAC CRC64;

Query Match 65.1%; Score 41; DB 2; Length 696;
Best Local Similarity 72.7%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFPA 11
Db 682 VPVLDEGLKA 692

RESULT 21
Q7WCJ7 PRELIMINARY; PRT; 696 AA.
AC Q7WCJ7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

```

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DN DNA ligase (EC 6.5.1.2).  
 GN Name=liga; Synonyms=dnaL, lig, lop, pdeC; OrderedLocusNames=BB3804;  
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=2287954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,  
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 Cardeno-Farraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 Rabinowitz E., Ruter S., Sanders M., Saunders D., Seeger K.,  
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 Unwin L., Whitehead S., Barrrell B.G., Masekell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 CC -!- FUNCTION: This protein catalyzes the formation of phosphodiester  
 CC linkages between 5'-phosphoryl and 3'-hydroxyl groups in double-  
 CC stranded DNA using NAD as a coenzyme and as the energy source for  
 CC the reaction. It is essential for DNA replication and repair of  
 CC damaged DNA (by similarity).  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + (deoxyribonucleotide)(N) +  
 CC {deoxyribonucleotide}(N+M) = AMP + nicotinamide nucleotide +  
 CC {deoxyribonucleotide}(N+M).  
 CC -!- SIMILARITY: Belongs to the NAD-dependent DNA ligase family.  
 DR EMBL; BX60448; CAE35778.1; -.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0003911; F:DNA ligase (NAD+) activity; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0006281; F:DNA repair; IEA.  
 DR GO; GO:0006260; F:DNA replication; IEA.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR001679; DNALigase.  
 DR InterPro; IPR004150; DNA\_ligase\_OB.  
 DR InterPro; IPR000445; HhH.  
 DR InterPro; IPR003583; HhH\_1.  
 DR InterPro; IPR004149; ZnF\_DNALigase\_C4.  
 DR Pfam; PF00533; BRCT; 1.  
 DR Pfam; PF01653; DNA\_ligase\_aden; 1.  
 DR Pfam; PF03120; DNA\_ligase\_OB; 1.  
 DR Pfam; PF03119; DNA\_ligase\_ZED; 1.  
 DR Pfam; PF00633; HhH; 3.  
 DR ProDom; P0003944; DNALigase; 1.  
 DR SMART; SM00292; BRCT; 1.  
 DR SMART; SM00278; HhH; 4.  
 DR SMART; SM00532; LIGANC; 1.  
 DR TIGRFAMs; TIGR00575; dnaL; 1.  
 DR PROSITE; PS0172; BRCT; 1.  
 DR PROSITE; PS01055; DNA\_LIGASE\_N1; 1.  
 DR PROSITE; PS01056; DNA\_LIGASE\_N2; 1.  
 KW Complete proteome; DNA repair; DNA replication; Ligase; NAD.  
 SQ SEQUENCE 696 AA; 75555 MW; CBB4177DB1837EA5 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 696;  
 Best Local Similarity 72.7%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFA 11  
 :|||||  
 Db 682 VPVDEDEGLKA 692

RESULT 22  
 Q83VW6

ID Q83VW6 PRELIMINARY; PRT; 702 AA.  
 AC Q83VW6;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Topoisomerase II subunit B.  
 DE Name=parYR;  
 OS Streptomyces roseochromogenes subsp. oscitans.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=149682;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DS 12.976;  
 RX MEDLINE=22491493; PubMed=12604514;  
 RA Schmutz E., Muehlenweg A., Li S.-M., Heide L.;  
 RT "Resistance genes of aminocoumarin producers: two type II  
 RT topoisomerase genes confer resistance against coumermycin A1 and  
 RT clorobiocin.";  
 RL Antimicrob. Agents Chemother. 47:869-877(2003).  
 CC -!- FUNCTION: DNA gyrase negatively supercoils closed circular double-  
 CC stranded DNA in an ATP-dependent manner and also catalyzes the  
 CC interconversion of other topological isomers of double-stranded  
 CC DNA rings, including catenanes and knotted rings (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
 CC of double-stranded DNA.  
 CC -!- SIMILARITY: Belongs to the type II topoisomerase family.  
 DR EMBL; AV136281; AA075469.1; -.  
 DR HSPB; P06982; IKZN  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.  
 DR GO; GO:0016853; F:isomerase activity; IEA.  
 DR GO; GO:0006304; P:DNA modification; IEA.  
 DR GO; GO:0006265; P:DNA topological change; IEA.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR002288; DNA\_gyrase\_C.  
 DR InterPro; IPR011558; DNA\_gyrase\_B.  
 DR InterPro; IPR001241; DNA\_topoisomII.  
 DR InterPro; IPR006171; Toprim\_dom.  
 DR Pfam; PF00204; DNA\_gyraseB; 1.  
 DR Pfam; PF00986; DNA\_gyraseB; 1.  
 DR Pfam; PF02518; HATPase\_C; 1.  
 DR Pfam; PF01751; Toprim; 1.  
 DR PRINTS; PR00418; TP12FAMILY.  
 DR ProDom; PD149633; DNA\_gyrase\_B; 1.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR SMART; SM00433; TOP2C; 1.  
 DR PROSITE; PS00177; TOPOISOMERASE\_II; UNKNOWN\_1.  
 KW ATP-binding; Isomerase; Topoisomerase.  
 SQ SEQUENCE 702 AA; 76721 MW; FA13366DA62D51E3 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 702;  
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFAP 12  
 :|||||  
 Db 286 VPVLDHGQWTP 297

RESULT 23  
 Q83WB8 PRELIMINARY; PRT; 702 AA.  
 AC Q83WB8;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE DNA gyrase subunit B.  
 DE Name=parYR;  
 OS Streptomyces rishiriensis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=68264;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 40489;  
 RC MEDLINE=22491493; PubMed=12604514;  
 RA Schmutz E., Muehlenweg A., Li S.-M., Heide L.;  
 RT "Resistance genes of aminocoumarin producers: two type II  
 RT topoisomerase genes confer resistance against coumermycin A1 and  
 RT chlorobiocin.";  
 RL Antimicrob. Agents Chemother. 47:869-877(2003).  
 CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-  
 CC stranded DNA in an ATP-dependent manner and also catalyzes the  
 CC interconversion of other topological isomers of double-stranded  
 CC DNA rings, including catenanes and knotted rings (By similarity).  
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
 CC of double-stranded DNA.  
 CC -1- SIMILARITY: Belongs to the type II topoisomerase family.  
 CC EMBL: AF205853; AA047226.1; -.  
 DR HSP; P06982; 1KZN.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.  
 DR GO: GO:0006304; P:DNA modification; IEA.  
 DR GO: GO:0006265; P:DNA topological change; IEA.  
 DR InterPro: IPR003594; ATPbind\_ATPase.  
 DR InterPro: IPR002288; DNA\_gyraseB\_C.  
 DR InterPro: IPR011558; DNA\_gyraseB.  
 DR InterPro: IPR001241; DNA\_topoisomII.  
 DR Pfam: PF00204; DNA\_gyraseB\_1.  
 DR Pfam: PF00986; DNA\_gyraseB\_C\_1.  
 DR Pfam: PF02518; HATPase\_C; 1.  
 DR PRINTS; PR00418; TPI2FAMILY.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00433; TOP2C; 1.  
 DR SMART; SM00177; TOPOISOMERASE\_II; UNKNOWN\_1.  
 DR PROSITE; PS00177; TOPOISOMERASE\_II; UNKNOWN\_1.  
 KW ATP-binding; Isomerase; Topoisomerase.  
 SQ SEQUENCE 702 AA; 76864 MW; D81B6665E50757D18 CRC64;  
 Query Match 65.1%; Score 41; DB 2; Length 702;  
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 IPVLDENGLFAP 12  
 :|||||:|  
 Db 285 VPVLDEHGQWTP 296  
 RESULT 24  
 ID O69998 PRELIMINARY; PRT; 707 AA.  
 AC O69998;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative DNA gyrase subunit B.  
 GN OrderedLocusNames=SC05822; ORFNames=SC5B8.12;  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;  
 RA Bentley S.D., Chater K.F., Cerdono-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J.R., Barrall B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RL Nature 417:141-147(2002).  
 CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-  
 CC stranded DNA in an ATP-dependent manner and also catalyzes the  
 CC interconversion of other topological isomers of double-stranded  
 CC DNA rings, including catenanes and knotted rings (By similarity).  
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
 CC of double-stranded DNA.  
 CC -1- SIMILARITY: Belongs to the type II topoisomerase family.  
 CC EMBL: AL393125; CAAL8520.1; -.  
 DR PIR; T35196; T35196.  
 DR HSP; P06982; 1AJ6.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.  
 DR GO: GO:0006304; P:DNA modification; IEA.  
 DR GO: GO:0006265; P:DNA topological change; IEA.  
 DR InterPro: IPR003594; ATPbind\_ATPase.  
 DR InterPro: IPR002288; DNA\_gyraseB\_C.  
 DR InterPro: IPR011558; DNA\_gyraseB.  
 DR InterPro: IPR001241; DNA\_topoisomII.  
 DR InterPro: IPR006171; Toprim\_dom.  
 DR Pfam: PF00204; DNA\_gyraseB\_1.  
 DR Pfam: PF00986; DNA\_gyraseB\_C\_1.  
 DR Pfam: PF02518; HATPase\_C; 1.  
 DR Pfam: PF01751; Toprim; 1.  
 DR PRINTS; PR00418; TPI2FAMILY.  
 DR PRODOM; PD149633; DNA\_gyraseB; 1.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00433; TOP2C; 1.  
 DR PROSITE; PS00177; TOPOISOMERASE\_II; UNKNOWN\_1.  
 KW ATP-binding; Complete proteome; Isomerase; Topoisomerase.  
 SQ SEQUENCE 707 AA; 77269 MW; 1A5427EB8EFB6661 CRC64;  
 Query Match 65.1%; Score 41; DB 2; Length 707;  
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 IPVLDENGLFAP 12  
 :|||||:|  
 Db 290 VPVLDEHGQWTP 301  
 RESULT 25  
 Q6CPB9 PRELIMINARY; PRT; 778 AA.  
 ID Q6CPB9  
 AC Q6CPB9;  
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Strain NRRL Y-1140 chromosome E of strain NRRL Y-1140 of Kluyveromyces  
 DE lactis.  
 GN ORFNames=KLLA0E06028g;  
 OS Kluyveromyces lactis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
 OX NCBI\_TaxID=28985;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RG GENOLEVURES;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boisrame A., Boyer J., Cattolico L., Confanieri F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,

RA Nicaud J.M., Nikoleki M., Oztas S., Ozier-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof B., With B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.,  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: CR382125; CAG99307.1; -.  
 SQ SEQUENCE 778 AA; 87407 MW; ACA0744901A55CE0 CRC64;

Query Match 65.1%; Score 41; DB 2; Length 778;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVLDENGL 9  
 DB 83 PILDENGL 90  
 :|||||

RESULT 26  
 Q71TF8  
 ID Q71TF8 PRELIMINARY; PRT; 2255 AA.  
 AC Q71TF8;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE DarB.  
 GN Name=darB;  
 OS Bacteriophage P1.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
 OC P1-like viruses.  
 OX NCBI\_TaxID=10678;  
 RN [1]\_TaxID=10678;  
 RP SEQUENCE FROM N.A.  
 RA Lobočka M.B., Rose D.J., Rusin M., Plunkett G. III, Samoedny A.,  
 RA Lehnher H., Yarmolinsky M.B., Blattner F.R.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=mod749::IS5 cl.100 mutant;  
 RA Lobočka M.B., Rose D.J., Rusin M., Plunkett G. III, Samoedny A.,  
 RA Lehnher H., Yarmolinsky M.B., Blattner F.R.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Lobočka M.B.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Lobočka M.B., Rose D.J., Rusin M., Plunkett G. III, Samoedny A.,  
 RA Lehnher H., Yarmolinsky M.B., Blattner F.R.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Rusin M., Samoedny A.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF234173; AAQ14093.1; -.  
 DR EMBL; AF234172; AAQ13985.1; -.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase C.  
 DR InterPro; IPR002296; N12N6\_mtfase.  
 DR Pfam; PF00271; Helicase C; 1.  
 DR PRINTS; PR00507; N12N6MTFRASE.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 SQ SEQUENCE 2255 AA; 251540 MW; 383BAA4D9D28341F CRC64;

Query Match 65.1%; Score 41; DB 2; Length 2255;  
 Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IPVLDENG 8  
 DB 925 VPVLDENG 932  
 :|||||

RESULT 27  
 AAQ13985  
 ID AAQ13985 PRELIMINARY; PRT; 2255 AA.  
 AC AAQ13985;  
 DT 03-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 03-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 03-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE DarB.  
 GN DarB.  
 OS Bacteriophage P1.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
 OC P1-like viruses.  
 OX NCBI\_TaxID=10678;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=mod749::IS5 cl.100 mutant;  
 RA Lobočka M.B., Rose D.J., Rusin M., Plunkett G. III, Samoedny A.,  
 RA Lehnher H., Yarmolinsky M.B., Blattner F.R.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=mod749::IS5 cl.100 mutant;  
 RA Rusin M., Samoedny A.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF234172; AAQ13985.1; -.  
 SQ SEQUENCE 2255 AA; 251540 MW; 383BAA4D9D28341F CRC64;

Query Match 65.1%; Score 41; DB 2; Length 2255;  
 Best Local Similarity 87.5%; Pred. No. 4.3e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVLDENG 8  
 DB 925 VPVLDENG 932  
 :|||||

RESULT 28  
 AAQ14093  
 ID AAQ14093 PRELIMINARY; PRT; 2255 AA.  
 AC AAQ14093;  
 DT 04-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 04-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 04-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE DarB.  
 GN DarB.  
 OS Bacteriophage P1.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
 OC P1-like viruses.  
 OX NCBI\_TaxID=10678;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=mod1902::IS5 cl.100 rev dmt(del)MB mutant;  
 RA Lobočka M.B., Rose D.J., Rusin M., Plunkett G. III, Samoedny A.,  
 RA Lehnher H., Yarmolinsky M.B., Blattner F.R.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF234173; AAQ14093.1; -.  
 SQ SEQUENCE 2255 AA; 251541 MW; 383BAA4D9D28341F CRC64;

Query Match 65.1%; Score 41; DB 2; Length 2255;  
 Best Local Similarity 87.5%; Pred. No. 4.3e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPVLDENG 8  
 :|||||

```

Db          925 VPVLDENG 932

RESULT 29
Q8RGZ3      PRELIMINARY;      PRT; 2462 AA.
ID Q8RGZ3
AC Q8RGZ3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hemolysin.
GN OrderedLocusNames=FN0132;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11899109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kypides N.C., Overbeek R.;
RA "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
RL EMBL; AE010527; AAL94338.1; -.
DR InterPro; IPR008619; Fil_haemagg.
DR InterPro; IPR006626; PbhI.
DR Pfam; PF05594; Fil_haemagg; 6.
DR SMART; SM00710; PbhI; 6.
KW Complete proteome.
SQ SEQUENCE 2462 AA; 267525 MW; 0DE3BF6FAFA5A26 CRC64;

Query Match          65.1%; Score 41; DB 2; Length 2462;
Best Local Similarity 87.5%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPVLDENG 8
      :|||||
Db      945 VPVLDENG 952

RESULT 30
Q8RI19      PRELIMINARY;      PRT; 2806 AA.
ID Q8RI19
AC Q8RI19;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hemolysin.
GN OrderedLocusNames=FN1817;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11899109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kypides N.C., Overbeek R.;
RA "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
RL EMBL; AE010484; AAL93916.1; -.
DR InterPro; IPR008619; Fil_haemagg.
DR InterPro; IPR008638; Haemagg_act_N.

```

```

DR InterPro; IPR006626; PbhI.
DR Pfam; PF05594; Fil_haemagg; 6.
DR Pfam; PF05860; Haemagg_act; 1.
DR SMART; SM00710; PbhI; 7.
DR TIGRFAMs; TIGR01901; adhes_NPXG; 1.
KW Complete proteome.
SQ SEQUENCE 2806 AA; 304992 MW; 314255277612B236 CRC64;

Query Match          65.1%; Score 41; DB 2; Length 2806;
Best Local Similarity 87.5%; Pred. No. 5.4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPVLDENG 8
      :|||||
Db      1128 VPVLDENG 1135

RESULT 31
Q7P6S5      PRELIMINARY;      PRT; 3119 AA.
ID Q7P6S5
AC Q7P6S5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hemolysin.
GN Name=FNVI407;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49256;
RA Kapatral V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haselkorn R., Overbeek R., Kypides N.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABF01000028; EAA24506.1; -.
DR InterPro; IPR008619; Fil_haemagg.
DR InterPro; IPR008638; Haemagg_act_N.
DR Pfam; PF05594; Fil_haemagg; 11.
DR Pfam; PF05860; Haemagg_act; 1.
DR TIGRFAMs; TIGR01901; adhes_NPXG; 1.
SQ SEQUENCE 3119 AA; 342583 MW; C03492B2DAE0CA50 CRC64;

Query Match          65.1%; Score 41; DB 2; Length 3119;
Best Local Similarity 87.5%; Pred. No. 6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPVLDENG 8
      :|||||
Db      1276 VPVLDENG 1283

RESULT 32
Q26740      PRELIMINARY;      PRT; 157 AA.
ID Q26740
AC Q26740;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conserved protein.
GN OrderedLocusNames=MTH644;
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;

```

```

RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaxH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AF000844; AB885149.1; -.
DR PIR; F69185; F69185.
DR InterPro; IPR000644; CBS.
DR Pfam; PF00571; CBS; 2.
DR SMART; SM00116; CBS; 2.
KW Complete proteome.
SQ SEQUENCE 157 AA; 17387 MW; 862166DC5B43E9A CRC64;

Query Match 63.5%; Score 40; DB 2; Length 157;
Best Local Similarity 63.6%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 IPIVLENGIIFA 11
Db 133 LPVIDENGRLA 143

RESULT 33
Q9V8G6 PRELIMINARY; PRT; 224 AA.
AC Q9V8G6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CG5164-PA (Glutathione S-transferase) (GH14654p)
GN Name=GetEl; Synonyms=Get3; ORFNames=CG5164;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Gabor G.L.,
RA Balow R.M., Baou A., Baxendale J., Bayraktaroglu L., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Beasley E.M.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Bueam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoatlin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy I., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,

```

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Testor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacieb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RX FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=20237668; PubMed=10773456;
RA Singh M., Silva E., Schulze S., Sinclair D.A.R., Fitzpatrick K.A.,
RA Honda B.M.;
RT "Cloning and characterization of a new theta-class glutathione-S-
RT transferase (GST) gene, gst-3, from Drosophila melanogaster.";
RL Gene 247:167-173(2000).
RN [8]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the GST superfamily.
DR EMBL; AF003799; AAF57701.1; -.
DR EMBL; AF179869; AAF64647.1; -.
DR EMBL; AY058383; AAL13612.1; -.

```

```
DR HSP; Q9GNE9; IJLV.
DR IntAct; Q9V8G6; -.
DR FlyBase; FBgn0034335; GstB1.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C-like.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
DR Pfam; PF02798; GST_N; 1.
DR Transferase.
SQ SEQUENCE 224 AA; 24959 MW; 148C2F359DBBC806 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 224;
Best Local Similarity 60.0%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDEGLF 10
Db :|||:|
57 VPMLDDNGTF 66

RESULT 34
Q7PSW6 PRELIMINARY; PRT; 233 AA.
AC Q7PSW6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000018720.
GN Name=ENSANGG0000016231;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the GST superfamily.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAB0100881; EAA04937.2; -.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
DR Pfam; PF02798; GST_N; 1.
SQ SEQUENCE 233 AA; 26064 MW; AB3222CAB0CC0FA3 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 233;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDEGLF 10
Db :|||:|
52 IPVLDDGDF 61

RESULT 35
Q8MUQ6 PRELIMINARY; PRT; 233 AA.
AC Q8MUQ6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glutathione S-transferase ul.
GN Name=Gstul;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=ZAN/U;
RX PubMed=12914673;
RA Ding Y., Ortelli F., Rossiter L.C., Hemingway J., Ranson H.;
RT "The Anopheles gambiae glutathione transferase supergene family:
RT annotation, phylogeny and expression profiles.";
RL BMC Genomics 4:35-35(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAN/U;
RA Ranson H.A., Ortelli F.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the GST superfamily.
DR EMBL; AF515521; AAW1888.1; -.
DR HSP; Q9GNE9; IJLV.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C like.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
DR Pfam; PF02798; GST_N; 1.
DR Transferase.
KW Transferase.
SQ SEQUENCE 233 AA; 26064 MW; AB3222CAB0CC0FA3 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 233;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPVLDEGLF 10
Db :|||:|
52 IPVLDDGDF 61

RESULT 36
Q6YUW2 PRELIMINARY; PRT; 236 AA.
AC Q6YUW2;
DT 03-JUL-2004 (TrEMBLrel. 27, Created)
DT 03-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNBa0078N11.32.
GN Name=OSJNBa0078N11.32;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005848; BAD16458.1; -.
KW Hypothetical protein.
SQ SEQUENCE 236 AA; 24733 MW; 44E41653230F7395 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 236;
Best Local Similarity 72.7%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLDEGLRAP 12
Db :|||:|
3 PFLDEQGLAAP 13

RESULT 37
BAD16458 PRELIMINARY; PRT; 236 AA.
AC BAD16458;
DT 23-APR-2004 (TrEMBLrel. 27, Created)
DT 23-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 23-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNBa0078N11.32.
GN OSJNBa0078N11.32.
```

```

OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
  clone:OSJNBa0078N11.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005848; BAD16458.1; -.
KW Hypothetical protein.
SQ SEQUENCE 236 AA; 24733 MW; 44E41653230F7395 CRC64;

  Query Match      63.5%; Score 40; DB 2; Length 236;
  Best Local Similarity 72.7%; Pred. No. 60;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVLDEGLFAP 12
Db      : |||||
        3 PFLDEQGLAAP 13

RESULT 38
Q841U3 ID Q841U3 PRELIMINARY; PRT; 333 AA.
AC Q841U3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Laci-1.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29307;
RA Baek C.H., Kim I.H., Kim K.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH laci-type DNA-binding domain.
DR EMBL; AY252118; AAP04598.1; -.
DR HSSP; P03023; 1LCC.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000843; HTH_Laci.
DR InterPro; IPR010982; Lambda like DNA.
DR InterPro; IPR001761; PeriplaBP/Laci.
DR Pfam; PF00356; Laci; 1.
DR Pfam; PF00532; Peripla_BP_1; 1.
DR PRINTS; PR00036; HTHLACI_-.
DR SMART; SM00354; HTH_LACI_1.
DR PROSITE; PS00356; HTH_LACI_1; 1.
DR PROSITE; PS00932; HTH_LACI_2; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 333 AA; 36097 MW; BF1E151EB5A77FFB CRC64;

  Query Match      63.5%; Score 40; DB 2; Length 333;
  Best Local Similarity 66.7%; Pred. No. 87;
  Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 IPVLDENGLFAP 12
Db      : |||||
        253 LSVLDENGLKVP 264

RESULT 39
Q7MI83 ID Q7MI83 PRELIMINARY; PRT; 333 AA.
AC Q7MI83;

```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcriptional regulator.
GN Name=V2633;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
  Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
  Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
  pathogen.";
RL Genome Res. 13:2577-2587(2003).
CC -!- SIMILARITY: Contains 1 HTH laci-type DNA-binding domain.
DR EMBL; AP005340; BAC95398.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000843; HTH_Laci.
DR InterPro; IPR001761; PeriplaBP/Laci.
DR Pfam; PF00356; Laci; 1.
DR Pfam; PF00532; Peripla_BP_1; 1.
DR PRINTS; PR00036; HTHLACI_-.
DR PROSITE; PS00356; HTH_LACI_1; 1.
DR PROSITE; PS00932; HTH_LACI_2; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 333 AA; 36096 MW; 8BD8D51793677FFB CRC64;

  Query Match      63.5%; Score 40; DB 2; Length 333;
  Best Local Similarity 66.7%; Pred. No. 87;
  Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 IPVLDENGLFAP 12
Db      : |||||
        253 LSVLDENGLKVP 264

RESULT 40
Q8D8N6 ID Q8D8N6 PRELIMINARY; PRT; 333 AA.
AC Q8D8N6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcriptional regulator.
GN OrderedLocusNames=VW11775;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
  Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH laci-type DNA-binding domain.
DR EMBL; AE016802; AAO10186.1; -.
DR HSSP; P03023; 1LCC.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000843; HTH_Laci.
DR InterPro; IPR010982; Lambda like DNA.
DR InterPro; IPR001761; PeriplaBP/Laci.
DR Pfam; PF00356; Laci; 1.
DR Pfam; PF00532; Peripla_BP_1; 1.

```

DR PRINTS; PR00036; HTH\_LACI.  
 DR SMARI; SM00354; HTH\_LACI; 1.  
 DR PROSITE; PS00356; HTH\_LACI; 1.  
 DR PROSITE; PSS0932; HTH\_LACI; 2; 1.  
 KW Complete proteome; DNA-binding; Transcription regulation.  
 SQ SEQUENCE 333 AA; 36096 MW; 21D720F8696765F0 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 333;  
 Best Local Similarity 66.7%; Pred. No. 87;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IPVLDEGLFAP 12  
 Db 253 LSVLDENGLKVP 264

Search completed: January 7, 2005, 10:03:56  
 Job time : 111 secs

This Page Blank (usp...)